



Review

Cite this article: Sitaram R, Sanchez-Corzo A, Vargas G, Cortese A, El-Deredy W, Jackson A, Fetz E. 2024 Mechanisms of brain self-regulation: psychological factors, mechanistic models and neural substrates. *Phil. Trans. R. Soc. B* **379**: 20230093.

<https://doi.org/10.1098/rstb.2023.0093>

Received: 29 September 2023

Accepted: 26 June 2024

One contribution of 16 to a theme issue 'Neurofeedback: new territories and neurocognitive mechanisms of endogenous neuromodulation'.

Subject Areas:

neuroscience

Keywords:

neurofeedback, brain–computer interface, brain–machine interface, active inference, reinforcement learning

Author for correspondence:

Ranganatha Sitaram

e-mail: ranganatha.sitaram@stjude.org

Mechanisms of brain self-regulation: psychological factors, mechanistic models and neural substrates

Ranganatha Sitaram¹, Andrea Sanchez-Corzo¹, Gabriela Vargas², Aurelio Cortese³, Wael El-Deredy^{4,5}, Andrew Jackson⁶ and Eberhard Fetz⁷

¹Multimodal Functional Brain Imaging and Neurorehabilitation Hub, Diagnostic Imaging Department, Saint Jude Children's Research Hospital, 262 Danny Thomas Place Memphis, TN 38105, USA

²Institute of Biological and Medical Engineering, Pontificia Universidad Católica de Chile, Diagonal Paraguay 362, Santiago de Chile 8330074, Chile

³Department of Decoded Neurofeedback, ATR Computational Neuroscience Laboratories, Kyoto 619-0288, Japan

⁴Brain Dynamics Lab, Universidad de Valparaíso, Valparaíso, Chile

⁵ValGrAI: Valencian Graduate School and Research Network of Artificial Intelligence – University of Valencia, Spain, Spain

⁶Biosciences Institute, Newcastle University, Newcastle NE2 4HH, UK

⁷Department of Physiology and Biophysics, Washington National Primate Research Center, University of Washington, Seattle, WA, USA

RS, 0000-0002-8577-8035

While neurofeedback represents a promising tool for neuroscience and a brain self-regulation approach to psychological rehabilitation, the field faces several problems and challenges. Current research has shown great variability and even failure among human participants in learning to self-regulate target features of brain activity with neurofeedback. A better understanding of cognitive mechanisms, psychological factors and neural substrates underlying self-regulation might help improve neurofeedback's scientific and clinical practices. This article reviews the current understanding of the neural mechanisms of brain self-regulation by drawing on findings from human and animal studies in neurofeedback, brain–computer/machine interfaces and neuroprosthetics. In this article, we look closer at the following topics: cognitive processes and psychophysiological factors affecting self-regulation, theoretical models and neural substrates underlying self-regulation, and finally, we provide an outlook on the outstanding gaps in knowledge and technical challenges.

This article is part of the theme issue 'Neurofeedback: new territories and neurocognitive mechanisms of endogenous neuromodulation'.

1. Introduction

Neurofeedback (NF) is an experimental technique for closed-loop brain training for animals and human participants to self-manipulate their brain signals [1–7]. NF provides an explicit sensory indicator of a neurophysiological process to enable individuals to modulate activation levels in specific ways to observe its effect on brain function and behaviour. NF has been recognized as a powerful approach to understanding brain–behaviour relationships and has shown potential in basic neuroscientific investigations and clinical rehabilitation to restore and enhance brain function and alleviate cognitive deficits in patient populations [6–10]. Methodological advances and experimental successes of NF led to the rapid development of the brain–computer interfaces (BCIs) and brain–machine interfaces (BMIs), which are methods for enabling humans or animals to directly control external devices without the involvement of peripheral limbs through the learned control of

specific features of neural activity [1–3]. The name BCI is generally attributed to non-invasive approaches, based on electroencephalography (EEG), functional magnetic resonance imaging (fMRI) or functional near-infrared spectroscopy (fNIRS). The term BMI is applied to invasive systems developed with electrocorticography (ECoG) and implantable electrodes. Hence, despite differences in implementation technology, methodological nuances and end users, NF, BCIs and BMIs have common neuropsychological principles. Therefore, it is instructive to draw on literature from these approaches for understanding the mechanistic underpinnings of brain self-regulation.

While NF presents promise for clinical treatment, the field still faces several problems and challenges. Studies have shown a great deal of variability among participants in learning brain self-regulation with NF, and many participants fail to learn self-regulation [6–11]. Hypothesized changes in behaviour or symptom improvement do not always follow changes in brain activity due to NF training. Furthermore, the observed behavioural changes or disease improvements typically do not persist over time. The research community often attributes the above problems to the prevalent short training regimens of NF, and calls for more controlled studies and clinical trials, more specific instructions to the participants for carrying out mental imagery, better feedback and reward contingencies, and more sophisticated multimodal and computational tools and techniques [12,13]. Researchers recognize that the neural mechanisms underlying NF learning are only beginning to be understood. NF researchers have expressed the need to develop a theoretical understanding of NF [5,6,14].

The field currently contends with several open questions that could be answered by a better understanding of the mechanistic nature of brain self-regulation. Some of those questions are: What is the optimal NF protocol for learning, and how does it vary with the target brain region(s)? When should mental strategies be suggested and explicitly provided, and when not? What are the optimal ways of providing feedback and reward? How do feedback and reward affect motivation, attention and learning?

It is now well recognized that the lack of sufficient scientific understanding of brain self-regulation adds to the difficulty in overcoming the above problems. Greater insight into the mechanisms underlying NF will help the research community in interpreting the results of data analysis of NF studies, providing explanations for negative findings, designing new protocols, improving existing protocols, controlling the quality of experimental protocols and brain and behavioural data, managing risk to the participants and the study and generating new hypotheses [15].

The purpose of the present article is to review the extant understanding of the neural mechanisms of brain self-regulation by drawing on findings from studies focusing on NF, BCI/BMI and neuroprosthetics. In the following sections, we will delve into the cognitive processes and psychophysiological factors affecting self-regulation, theoretical models of self-regulation, neural substrates underlying self-regulation, and finally, provide an outlook on the outstanding gaps in knowledge and technical challenges.

2. Cognitive processes and psychological factors of brain self-regulation

Learning to control brain activity in a specific manner is influenced by cognitive processes modulated by psychological factors [16,17] (see figure 1).

(a) Cognitive processes

Self-regulation of brain activity through NF has been proposed to involve two distinct cognitive processes: discrimination and self-maintenance [16]. Discrimination is the ability to perceive and identify the neural variable associated with the feedback signal. Self-maintenance is the ability to affect/influence the neural variable and change it in the intended manner. The hypothesis is that, when the above two skills are acquired through NF training, they would allow the participant to regulate the neural variable through a volitional psychosomatic process [18]. Hence, NF training should be evaluated by assessing the acquisition of discrimination and self-maintenance as dual skills. Unfortunately, there is a shortage of empirical or theoretical studies to test the above hypothesis.

Human NF studies have used two distinct experimental paradigms, ‘explicit’ and ‘implicit’ [17]. During explicit NF, the participant observes the feedback signal as a representation of the biological signal that needs to be controlled during training. For example, the participant sees the changes in the haemodynamic activity of the posterior parietal cortex as changing bars of a thermometer, or a circle’s size [21]. In this approach, the participant knows he/she is participating in NF training. The user learns and gains control over the relationship between the neurobiological signal and the changing experimental conditions.

In the implicit NF paradigm, the participant may not directly perceive or be aware of the biological signal, or both. The NF signal modifies the experimental condition (for example, changing the visual stimulus), and the participant does not observe the biological signal in any direct form. Thus, the participant is unaware of undergoing NF training and that the brain signal somehow influences the experiment. To cite an example scenario, as part of an experiment, participants can be asked to play a video game whose parameters (level of difficulty, the number of points won, etc.) change according to the brain activity. Implicit feedback does not provide information about the biological signal to the participant but incorporates indirect interactivity with the participant, thereby influencing the control of brain activity [19,22,23].

Some NF approaches seem to fall somewhat in between the two paradigms, i.e. between implicit and explicit NF. Decoded neurofeedback (Decnef, [20]) is one good example, in that participants know an NF experiment is taking place and receive explicit feedback about the targeted brain activity on each trial; however, they do not know what the feedback and/or how the feedback is related to their task performance [20]. A recent review has clarified the debate by discussing NF in terms of overt and covert processes [24].

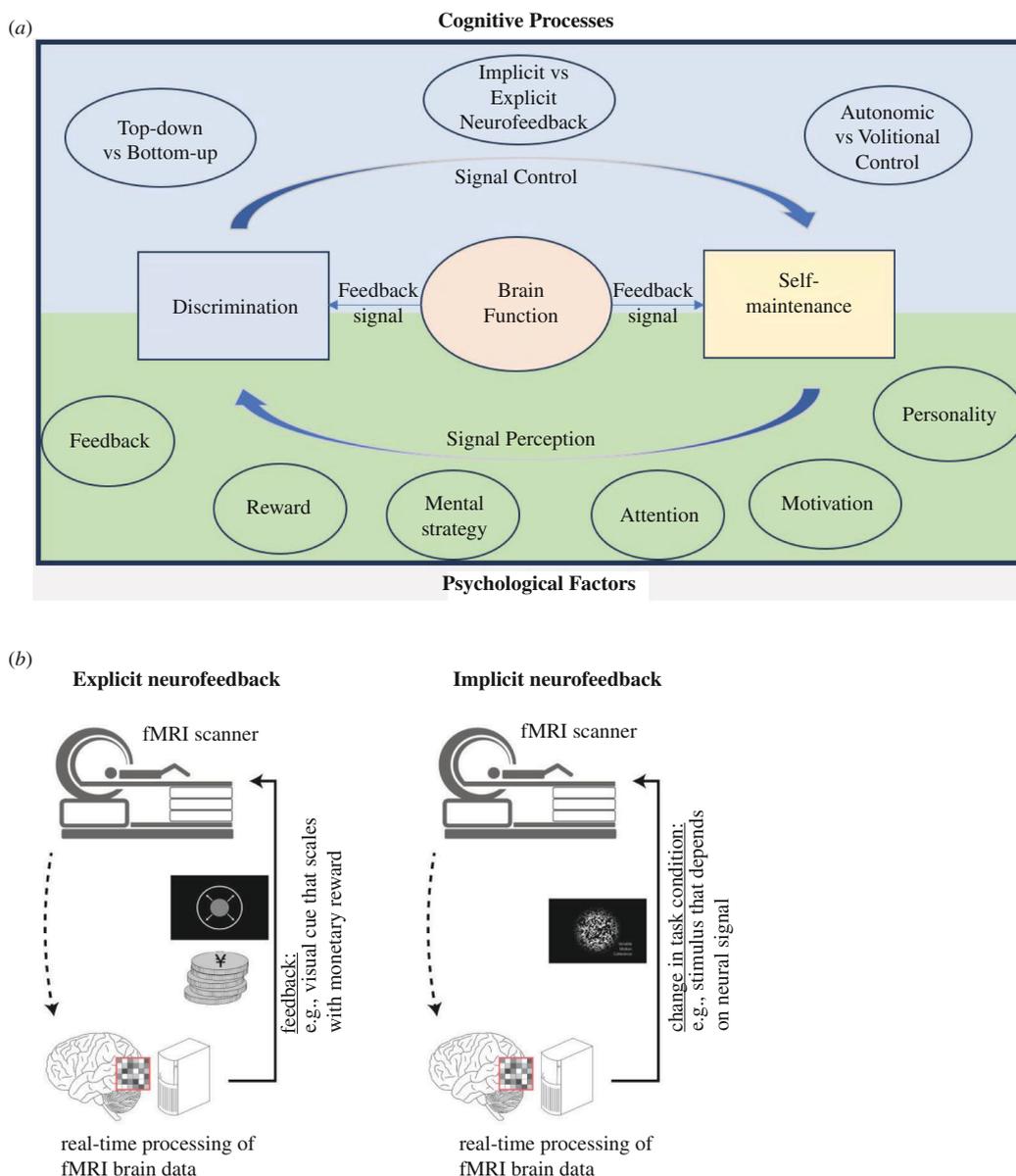


Figure 1. (a) Cognitive processes and psychological factors involved in brain self-regulation. Self-regulation of brain activity through neurofeedback (NF) has been proposed to involve two distinct cognitive processes: discrimination and self-maintenance [16]. Discrimination is the ability to perceive and identify the neural variable associated with the feedback signal. Self-maintenance is the ability to affect/influence the neural variable and change it in the intended manner. The hypothesis is that the above two skills acquired through NF training will allow the participant to regulate the neural variable through a volitional psychosomatic process [18]. Cognitive processes and psychological factors affecting NF have been identified in the literature. The most essential cognitive processes are related to top-down and bottom-up processing, implicit and explicit NF paradigms, and autonomic and volitional control. Psychological factors related to NF include feedback type, reward and mental strategy; participant attention and motivation levels; and personality-related features. (b) Explicit versus implicit NF. Human NF studies have used two distinct experimental paradigms, 'explicit' and 'implicit' [17]. During explicit NF, the participant observes the feedback signal as a representation of the biological signal that needs to be controlled during training. In this approach, the participants know they are participating in NF training. In the implicit NF paradigm, the participant does not directly perceive nor is aware of the biological signal. The NF signal modifies the experimental condition (for example, changing the visual stimulus) and the participant does not observe the biological signal in any direct form (modified from Renton *et al.* and Cortese *et al.* [19,20]).

There is no precise equivalent of the explicit/implicit distinction for animal NF paradigms, but studies vary in the extent to which animals are provided with feedback signals. Early animal NF studies delivered a reward, often associated with sensory cues, to train subjects to modulate brain signals. For example, the deflection of a meter arm could provide monkeys with a continuous signal showing how close their neural activity was to a reward threshold [25]. The meter deflection could signal complex combinations of neural and muscular activity, and as these combinations were changed, monkeys quickly changed their response patterns as required to obtain a reward [26]. Such NF could also signal neural patterns unlikely to have motor or sensory correlates, such as the relative number of epileptic bursts and normal firing of single cells in an epileptic focus [27].

More recently, the development of BMI paradigms prompted a renewed interest in NF. BMI-NF paradigms using non-human primates typically begin with subjects controlling a visual cursor with natural movements (e.g. reaching with the arm) while the activity of a large number of neurons is recorded [28]. This provides an initial training dataset to build a decoder that allows the cursor to be controlled directly by neural activity. In some experiments, this decoder has then been perturbed systematically, for example, by changing how some or all channels influence cursor movement requiring animals to learn a new control strategy for the BMI-NF [29,30]. Evidence suggests that the neural activity associated with natural movements may bias the exploration

of these new strategies. One method to analyse the population activity is to represent the instantaneous state of all the recorded neurons as a point in a high-dimensional neural 'state-space', where each dimension reflects the firing rate of a single neuron. Because the activity of individual neurons is often correlated with others, it is possible to extract lower dimensional subspaces or 'neural manifolds' within the full state-space that are associated with the movements in the training dataset. Evidence suggests that if BMI-NF control can be achieved with activity patterns that remain within these manifolds, it can be learned more quickly than a BMI-NF task that requires activity outside of the manifold [30]. One interpretation of this finding is that BMI-NF learning, at least within a single session, consists of a rapid reassociation between a pre-existing repertoire of activity patterns and the new goals of BMI-NF control [31]. Note that the subjects in these experiments typically would have been extensively trained with natural movement tasks prior to any BMI-NF protocols being introduced, and such pre-training may play a role similar to explicit instructions in human NF studies that establish an *a priori* expectation about strategies that might be successful. Alternatively, the neural correlations that constrain the neural manifolds may be imposed by the connectivity within the network of neurons. In any case, acquiring 'off-manifold' solutions typically requires several days of training [32]. Such multi-day learning is associated with an initial reduction in correlations between neurons (interpreted as an unstructured, trial-and-error exploration) followed by a progressive increase in correlated variability (interpreted as the formation of a new neural manifold appropriate for the BMI-NF) [33]. Similar learning processes may be required in those human NF tasks where an appropriate mental strategy is not provided in advance to subjects.

NF researchers have conflicting views on whether NF training leads to subconscious regulation of brain activity or represents volitional regulation brought about by the participant's intention. Some reports suggest that NF is more efficient when based on volitional self-regulation with consciously chosen mental strategies [34], while others maintain that regulation can be achieved without volitional control [7,35,36]. Except for a very few empirical studies [37], little has been done to directly test the above two hypotheses. Sepulveda and colleagues found through an empirical comparison of NF parameters that the feedback signal resulted in a significant learning effect. However, anecdotal reports by practitioners have expressed that feedback learning was hard to achieve until a cognitive strategy (e.g. imagery) was provided. The effectiveness of learning with or without volitional control probably depends on a variety of factors, including context, experimental paradigm, as well as the targeted brain state (e.g. the overall activity in the brain area or the pattern of activity reflecting some specific representation).

Another major question is whether brain self-regulation is a top-down or bottom-up neural mechanism. Top-down processing involves higher cognitive functions originating predominantly in the prefrontal cortices, leading to downstream processing in the sensorimotor and association areas [34,38]. Bottom-up processing originates in the primary and secondary sensory processing areas and drives prefrontal cortical brain activity [39]. These two views may lead to different ways of designing and conducting NF. Based on the bottom-up view, an experimenter may design NF training in terms of conditioning strategies, reinforcement schedules, shaping and chaining [23,40,41]. The top-down model may form the basis for instructing participants to use mental imagery to regulate activation in different brain areas.

To resolve the above contradiction, some researchers have proposed the dual-process mechanism of NF [5,42], according to which both the bottom-up and top-down mechanisms actively interact to support NF learning depending on the task requirements. The bottom-up, operant conditioning-related NF involves automatic and capacity-free processes, while the top-down, cognitively involved NF is a controlled and capacity-limited process. At the connectionist network level, the two processes could be seen as integrated in a network that combines the two separate top-down and bottom-up networks. In the following, we will elaborate on two extant models, namely, reinforcement learning and active inference, which represent two leading models roughly mapping onto the above two opposing approaches. In §3, we will discuss theoretical models of brain self-regulation based on rigorous mathematical formulation that enable researchers to conduct computational simulations and empirical tests to test the above two theories of brain self-regulation.

(b) Psychological factors

The NF literature suggests that the following experimental conditions and variables may influence NF learning and performance: feedback contingency, reward, instructed or self-developed mental strategies and duration of the training [6]. The major psychological variables in empirical literature include attention, motivation, mood and personality [43].

Contingency refers to the conditional probability of reinforcement, in terms of feedback and reward, given behavioural and neurophysiological responses or a failure to respond. The study of feedback and reward contingencies in the NF literature has been limited [37,44]. Investigation of feedback contingency would include studies on different modalities (visual, auditory, tactile, etc.) of the response-contingent stimuli, their different physical properties (such as amplitude, rate and complexity) and the different functional relationships between the response and the feedback [17].

NF learning is affected by the temporal contiguity of feedback and reward, which refers to the time interval between the neurophysiological response and the presentation of the reinforcement [45,46]. In the context of fMRI and fNIRS NF, the intrinsic delay (in seconds range) between the neural activity that is being regulated and the blood oxygenation-dependent (BOLD) signal changes due to the slow haemodynamic response is a critical factor to be considered, too. In comparison to the BOLD response, the scalp EEG has excellent temporal resolution with millisecond precision, allowing for a higher signal-to-noise ratio in the feedback information due to more samples of the data per time point [47]. In general, the onset of the haemodynamic response, measured by fMRI [48–50] and fNIRS [51,52], lags neural activity in the seconds range (the BOLD response, typically, starts 1 s after neural activity, takes 6 s to peak and takes 11 s to return to baseline from the onset of the haemodynamic response [53]). However, recent studies have shown that the BOLD response to neuronal activity can be more nonlinear and rapid than is commonly held, suggesting that fMRI can reflect fast neuronal dynamics [54]. The time involved in the acquisition and computation of the feedback signal would add to the delay in the feedback, affecting its temporal

contiguity. Whether such delayed temporal contiguity degrades or improves self-regulation learning is an open question [6]. A study showed that intermittent feedback (about 20 s delay) is more effective than continuous feedback presentation when the participant uses mental imagery for self-regulation [44]. The argument explains that intermittent feedback could be more advantageous in certain situations, such as during the early stage of learning because it does not interfere with ongoing mental imagery and associated dual-task cognitive overload. Whether intermittent or continuous feedback is more effective in explicit and implicit NF studies is still to be investigated. Some concrete indications come from simulation work, where cognitive strategies responded better to continuous feedback while intermittent feedback instead was more effective for automatic reward-based learning in implicit NF [35].

In animal studies operantly conditioning firing rates of neurons in the motor cortex, the NF signals can be instantaneous, providing signals that even precede any associated movements, which typically occur approximately 100–200 ms later. Often, feedback is multi-dimensional, like cursor control in two or three dimensions. Therefore, feedback signals can be highly informative in spatial and temporal domains. In this respect, at least, BMI-NF perturbation experiments share similarities with traditional motor learning paradigms. For example, the perturbation used by Jarosiewicz *et al.* [29], in which the mapping of a subset of neurons to cursor movements was rotated, generated errors like classic visuomotor rotation perturbations (in which the mapping from arm movement to the cursor is rotated) [29]. This may explain why the predominant strategy seen in response was a global re-aiming to rotated targets. Motor learning involves multiple processes occurring over different time scales in various brain areas [55], with a fast error-based learning system probably involving the cerebellum [56]. This rapid recalibration of the sensorimotor map may be equivalent to the early reassociation within the neural manifold in BMI-NF paradigms. By contrast, a slower-to-learn but also slower-to-forget reinforcement learning system responsible for skill learning preferentially operates when only low-dimensional ‘success-based’ feedback is available [57], probably dependent on long-term potentiation-like plasticity mechanisms involving the motor cortex [58]. In the context of BMI-NF paradigms, this slower skill-learning system may underlie the emergence of new neural manifolds over several training days. Evidence in rodents suggests that such NF learning may rely on corticostriatal plasticity [59] and can be prevented by optogenetic inhibition of the basal ganglia [60]. Moreover, operant conditioning of neural activity in freely behaving monkeys can be achieved with feedback mediated through stimulation of reward centres alone [61]. By this reasoning, human NF studies, especially those based on haemodynamic signals with significant time delays mediated by low-dimensional ‘success-based’ reinforcement rather than informative error signals, might similarly be expected to depend on corticostriatal skill-learning mechanisms. Interestingly, the consolidation of such long-term skill learning depends on sleep and, in particular, the reactivation in sleep of patterns of neural activity associated with the performance of the skill when awake [62]. This reactivation process involves the repeated firing of neural circuits that were initially used when learning the task, thereby strengthening the synapses that encode the learning. Studies in rodents have similarly shown that neural ensembles involved in BMI-NF task learning are reactivated in subsequent sleep [63]. However, the role of sleep-dependent consolidation processes in human NF studies has been surprisingly neglected.

Two major procedural elements have been used in extant studies for training voluntary self-regulation in humans: instructions and response-contingent stimulation [5,17]. Animal studies of NF and BMIs have strictly relied on feedback and reward following the operant learning paradigms [25]. While experimental work on voluntary control has tended to focus more on the role of feedback, older NF studies observed that instructions are not neutral in influencing voluntary control [64]. Hence, the investigation of experimental instructions is essential for the analysis of voluntary control. Furthermore, more recent developments in BCIs and real-time fMRI NF have emphasized designing experiments incorporating sensitive yet robust measures that control placebo effects and instruction with control conditions [65]. While providing explicit instructions to human participants to use mental imagery and monetary reward is quite common in recent work on NF to enhance and accelerate learning, the optimal strategy for improving volitional control remains unclear.

During the early period of biofeedback development, sophisticated procedures of operant training, such as shaping and chaining, received attention in the literature [66]. Such approaches have the potential to improve learning brain self-regulation. However, more recent work on human NF in EEG and fMRI studies has not explicitly investigated them.

Sepulveda *et al.* [37] investigated the differential effect of feedback, explicit instructions and monetary reward on healthy individuals to up-regulate the BOLD signal in the supplementary motor area (SMA) using fMRI NF. Four groups were trained in a two-day protocol: (i) group F with feedback only, (ii) group FR with feedback and monetary reward, (iii) group FI with feedback and instruction for carrying out motor imagery, and (iv) group FIR with feedback, monetary reward and instruction. Their results showed that group F participants learned to increase the BOLD signal in the SMA significantly from day 1 to day 2, while the other groups did not show a significant learned increase. Additionally, group FR attained the highest BOLD signal amplitude in SMA during the training, although the rise of the SMA signal was not significant from day 1 to 2. The two groups instructed to use motor imagery did not show a significant learning effect over the two days [37]. It is possible that the learning curves may be quite different across manipulations, i.e. a fast-learning curve quickly reaching a plateau for group FR and/or FIR (thus not showing a significant increase from day 1 to 2) as opposed to a slow-learning curve in group F (thus showing a significant increase from day 1 to 2). It is unclear whether these effects are specific to the SMA and the type of mental imagery instructed in the above study. More studies in different functional regions and brain networks are required to consolidate our understanding of the effect of experimental factors on self-regulation.

A major challenge in human NF studies is to find an effective reward as extrinsic motivation for human participants (i.e. food, money, gifts, encouraging words). Extrinsic rewards, however, can often have a negative effect on performance. Lepper *et al.* [67] showed through their psychological experiments that extrinsic reward may act as short-term stimulation but induce longer term depression of motivation [67]. How intrinsic reward should be identified and measured online and reinforcement provided accordingly are still open questions. A working hypothesis is that NF training is expected to develop the above two primary cognitive skills for successful regulation: discrimination and self-maintenance [17]. Accordingly, if NF training

leads to improvements in the above two skills, intrinsic motivation would be rewarded. Identifying cognitive strategies and experimental parameters to improve those skills may be more effective as rewarding information than extrinsic rewards such as money or food. Providing instructions to participants for using mental strategies and imagery could lead to enhancing intrinsic motivation.

Recently, Kadoush & Staunton conducted a systematic review of the NF literature to evaluate the psychological factors that influence NF learning carefully [68]. Their findings showed that attentional variables are significant to both performance and education; motivational factors and mood were moderate predictors of success. The authors suggest that future research should systematically manipulate psychological variables such as motivation or mood. Furthermore, non-responders could be tested with different NF parameters to understand whether their non-response is specific or general [68]. Such research will help develop better approaches for non-responders and improve the efficacy of NF.

3. Models of brain self-regulation

In this section, we will consider two major models of brain self-regulation, namely, reinforcement learning and active inference, which are two rigorously formalized models for analysis and comparison.

(a) The reinforcement learning model

Reinforcement learning (RL) is a general computational/modelling framework concerned with how agents learn to select actions that will maximize future cumulative rewards. In its original form, RL was discussed primarily to explain associative learning by which an association between two stimuli or between a behaviour and a stimulus is learned. Associative learning states that the probability of a physiological response is increased when a reinforcing stimulus follows that response. The theory focuses on three parameters, namely, discriminative stimulus, responses and reinforcers. When the response is reinforced in the presence of a discriminative stimulus and no other stimuli, the increase in response probability will occur only in the presence of that stimulus. In the case of an NF experiment, a reinforcing stimulus could be the feedback of the brain activity, for example, in the form of an increase in the bars of a thermometer in proportion to the amplitude of the BOLD signal in a brain region relative to baseline or a given reference activity.

Two different mechanisms of RL have been proposed in the literature [17,69,70]. Model-based RL is goal directed and based on the participant's internal model, e.g. in explicit NF, the participant has a goal to regulate the feedback signal in a specified direction [71,72]. In the model-free RL, the participant has no model of the environmental events and learning happens by the simple association of stimulus and response, e.g. in implicit NF, the participant does not receive information about the feedback signal directly [71].

Certain critical components of RL, like the choices made by participants regarding behaviour and exploration while learning, the predictions they form about the results of their choices and the ensuing prediction errors, are often overlooked or not recorded in practice and not included in NF analysis. As a result, the significance of these underlying parameters in NF learning has not been fully and systematically explored. It is necessary to recognize the two most critical dimensions in NF that define structurally different learning regimes to establish the correct interaction of the RL elements applied in NF. The first dimension involves the goals of regulation, which are mainly divided into explicit and implicit (although intermediate protocols exist). The second dimension is the feedback presentation time that separates continuous protocols (immediate feedback) from intermittent ones (delayed feedback or separated from the regulation phase).

Lubianiker *et al.* [39] proposed a general framework of NF based on RL to understand and evaluate the parameters that guide the learning of neural regulation. They suggested that this framework can generate new proposals to enhance the effectiveness of NF protocols [39]. In their model of RL, agents learn which actions they should take, guided by environmental reinforcement. In their framework, RL models the environment with three elements: states, actions and rewards. In NF protocols, actions change the state of the environment, and agents learn a policy that maps states to actions to maximize long-term rewards. Rewards depend on states and actions. Lubianiker *et al.* [39] formally show how RL elements, such as value functions, policies and credit assignment problems, are represented in different NF protocols. We present below a summary of the formal elements, but please refer to Lubianiker *et al.* [39] for a more in-depth understanding.

States at a specific trial n and point in time t , i.e. $x_n(t)$, depend on the previous state $x_n(t-1)$ and the action taken at that time $A_n(t-1)$, given by

$$x_n(t+1) = f[x_n(t), A_n(t)] + \varepsilon_n^x(t), \text{ where } \varepsilon_n^x(t) \text{ is the state noise.}$$

Actions A , understood as internal or external bodily movements, are a fundamental component in RL problems. In NF, actions are purely mental and hence neuronal events. An action at a specific trial n and point in time t , $A_n(t)$ will depend on the overall intentional strategy A_n and the subconscious strategy $a_n(t)$, given by

$$A_n(t) = A_n + a_n(t) + \varepsilon_n^a(t), \text{ where } \varepsilon_n^a(t) \text{ is the action-related noise.}$$

Rewards are the third central element in RL. Rewards can be immediate, associated with a specific action and state, or long term, determining the success of a response. The immediate reward $r_n(t)$ is a function of the state, given by

$$r_n(t) = g[x_n(t)] + \varepsilon_n^r(t), \text{ where } \varepsilon_n^r(t) \text{ is the reward noise.}$$

In explicit protocols, actions are usually mental strategies that participants explore, whose space may be limited or expanded according to task instructions. In explicit protocols with intermittent feedback, states are separated for regulation periods x_{regulate} and for reward presentation and evaluation of action utilities x_{feedback} . Intermittent feedback faces the challenge of

temporal credit assignment, which is the delay between the neural activity event and the subsequent feedback. Temporal credit assignment involves determining how credit should be assigned to intermediate actions in a sequence of decisions. However, intermediate feedback has the advantage of separating regulation and feedback states [73,74]. On the other hand, explicit protocols with continuous feedback do not have the temporal credit assignment problem because the feedback is provided continuously, but the participant's attention could be split between conscious action and evaluation of its utility since we have a combined state $x_{\text{regulate} + \text{feedback}}$.

A key question is what determines the 'state' in RL? Based on previous work on neural representations, and the spectrum of activity patterns that can be elicited at any given time by a population of neurons, it is evident that neural systems are much more constrained than they would superficially appear [31,75]. This means that, even in the context of fMRI signals, the mapping between fMRI multivoxel patterns and the underlying neural activity patterns are hypothesized to be closely matching [7,36].

What evidence do we have that RF is a plausible mechanism underpinning NF learning? Simulation work has shown activity patterns in artificial neural networks can be reinforced via a simple learning rule that updates the weighted connections between units in a manner compatible with standard decoder-based NF experiments, by computing a feedback signal from smoothed and delayed neural activity decoding (akin to fMRI, [36]). Similarly, Oblak *et al.* [76] demonstrated through simulations that intermittent feedback of decoded neural activity can be rapidly learned automatically, i.e. consistent with a basic RL mechanism [76]. A re-analysis of published work, particularly of decoded NF experiments, has highlighted how feedback elicits commensurate activity in the basal ganglia and medial prefrontal cortex [36], areas typically involved in RL processing [77–79]. Interestingly, the orbitofrontal cortex shows activity tracking cumulative failures in NF training [80]. Hence, learning to self-regulate brain activity is consistent with learning goal-directed actions, as in model-based RL. Finally, a recent study found the grey matter volume of the right putamen—a key region in associative/instrumental learning—to be predictive of NF success [81].

Animal literature is richer in studies that have explicitly linked BCI/BMI control of neural activity to RL mechanisms. In a foundational study, Koralek *et al.* [59] demonstrated that learning intentional neuroprosthetic skills depends on corticostriatal circuits and striatal synapses receiving cortical inputs (e.g. striatum and interactions between the striatum and motor cortex). The brain appears to learn faster to control new activity patterns that fall within existing neural population structures, as opposed to entirely new activity patterns, inconsistent with existing neural network structures [30]. The rapid co-optation of a pre-existing structure to achieve new behavioural outcomes resembles model-based RL, whereas the longer term development of de novo neural strategies based on unguided exploration may be better described by model-free RL. An interesting hypothesis postulates that RL operates directly on neural dynamics [82]. RL provides a simple rule to change neural firing rates, which results in altered dynamics. This mechanism can also provide a solution to the issue of how the brain can re-enter a certain state (as a trajectory in neural state space) and refine over time cortical dynamics to achieve target states more reliably [82]. Given that RL seems widely present in neural computations related to learning and behaviour changes (e.g. even visual perceptual learning can be accounted for by RL, see [83]), RL presents itself as a plausible model for explaining NF learning.

(b) The active inference model

Active inference is a theoretical framework for understanding learning and decision-making in the brain, based on the free energy principle [84–88]. It posits that organisms constantly minimize surprise or uncertainty in their environment by making predictions and actions. This is achieved by the brain constructing internal representations (models) of the environment and generating predictions about incoming sensory information. When the internally generated predictions do not match reality, a prediction error arises, requiring the brain to update its internal models, to improve future predictions, thus reducing future surprise or uncertainty. During perception, the brain anticipates sensory input, allowing it to actively sample information that aligns with its predictions, thus minimizing discrepancies. During action, the brain selects behaviours that minimize the discrepancy between predicted and actual outcomes, leading to adaptive responses. This process drives learning by encouraging the selection of actions that align with desired outcomes. By viewing NF learning as a process of minimizing uncertainty, active inference may contribute to our understanding of how brain self-regulation becomes possible.

Constructing internal models entails inferring the hidden (unobservable) states of the world that cause the observations, and optimizing actions to update the models or to adapt to change. It is suggested that this is carried out by using Bayes' rule in a variational scheme [86]: the brain generates a probability distribution of the sensory information (p) and optimizes it using an approximate distribution (q). The optimization borrows the concept of 'free-energy' from statistical physics, where free energy represents an upper limit on the difference between (p) and the (q). Minimizing free energy corresponds to maximizing the brain's ability to predict and control its environment effectively.

During NF training, the generative model in the brain quantifies the relation between sensory observations from the feedback (o) and the brain's hidden states that caused them (s). This relation is captured by a likelihood mapping (A), which encodes beliefs about how the brain's hidden states are related to the observations that they generate (figure 2b). The generative model also accounts for the knowledge of the temporal dynamics of states in the form of beliefs about state transitions. This is captured by a transition matrix (B), which encodes the probability of being in some brain state s^1 at some time step $t + 1$, given that the system was in brain state s^2 at time step t . The control of the transitions is made by the policy (π), a belief over the sequence of mental actions (a), with $\pi = (a_1, a_2, \dots, a_T)$ in the time horizon $T : t_1, \dots, t_n$; where policy selection is implemented through the updating of beliefs about state transitions, informed by the consequences of mental action.

Two parameters influence policy selection: the prior preference (C) and the prior policy precision (γ). The parameter C defines the agent's pre-existing beliefs about sensory outcomes, irrespective of the policy's free energy. Essentially, it represents

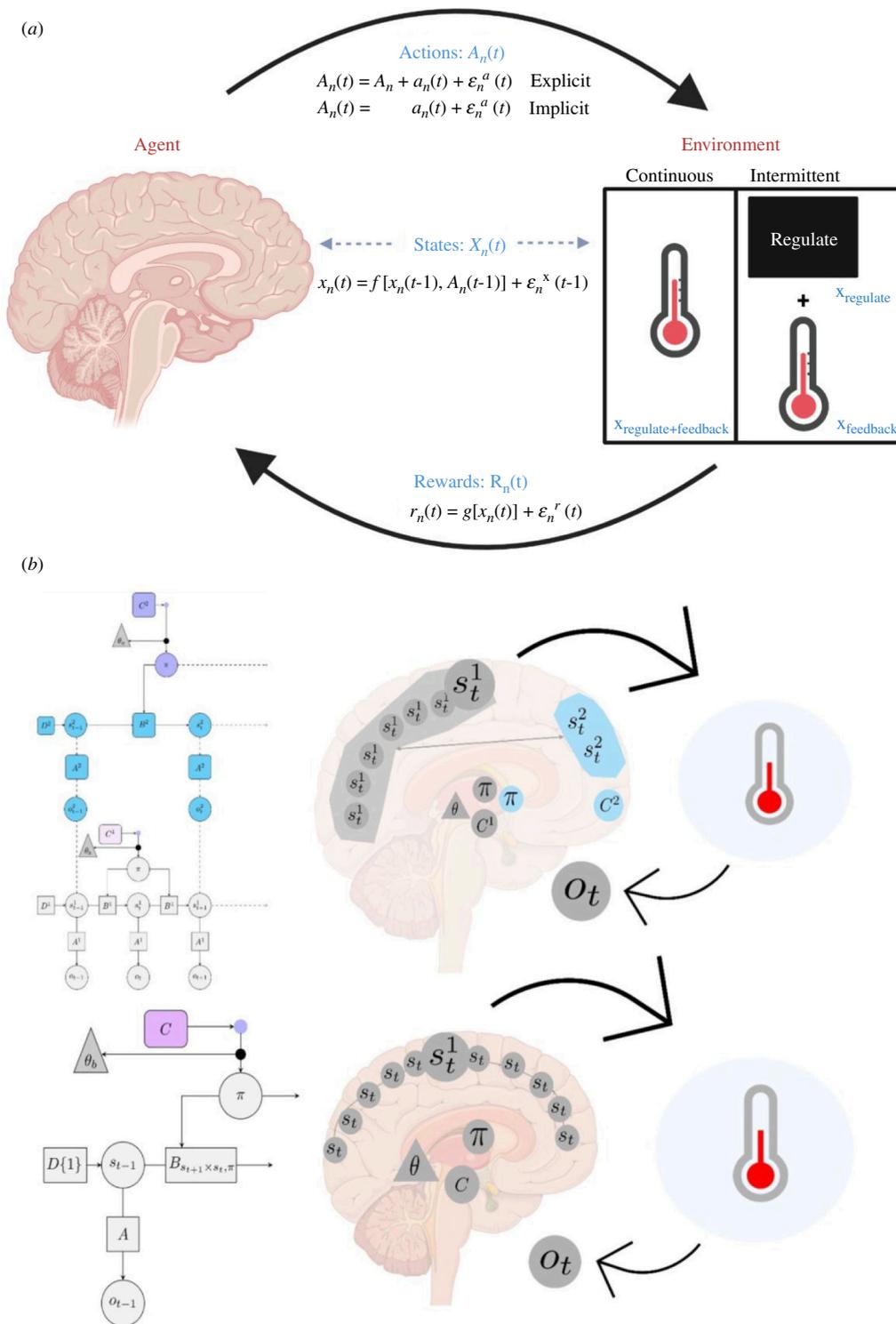


Figure 2. Reinforcement learning (RL) and active inference models. (a) RL elements in neurofeedback (NF). In the context of NF, the agent is the NF trainee, and the environment is the task interface that informs the agent about its states, identified as $x_n(t)$ at time t during trial n . These states correspond to changing cues within the task, such as 'rest' or 'regulate', along with observations that are assessed for rewards $r_n(t)$. The agent engages in cognitive actions, denoted as $A_n(t)$, which impact the transition between different neural states—essentially the brain's neural dynamics across various regions (also represented as $x_n(t)$). This activity, in turn, modifies the feedback received, creating a feedback loop. This feedback loop blurs the lines between the agent and the environment within NF frameworks, a phenomenon that intertwines their roles. Created in Biorender and adapted from Lubianiker *et al.* [39]. (b) Computational neuroanatomy of brain self-regulation by active inference. Computational architectures of active inference-based generative models, within a partially observable Markov decision process (POMDP) framework [89] are depicted, with hierarchical structure (top) and without hierarchical structure (bottom), in the context of an NF task. The mapping of model elements (left column) to anatomical brain areas (right column) aims to match information processing and the corresponding anatomical networks based on existing knowledge and theoretical assumptions regarding brain regions and their interactions [87,90–92]. This highlights the importance of hierarchical architecture in simulating meta-awareness and controlling mental actions, reflecting top-down information processing observed in brain data. The model uses notions of a POMDP [89], encompassing hidden states (s), observations (o), likelihood matrices (A), transition matrices (B), expected action sequences or policies (π), parameters capturing prior beliefs (C), and memory formation (θ). Following an NF experiment, the agent infers the hidden brain states (s) from NF observations (o) and the consequences of mental actions (π). Mental actions (π) transition brain-hidden states from the current state ' s ' at time step ($t + 1$) to a new state ' s ' according to the transition probability represented by matrix B . Sensory observations (o) from the new state are generated from a likelihood distribution given by A . The search is characterized by

Figure 2. (continued.)

a high volume of hidden states and narrow observations (such as a visual thermometer). In the one-layer structure, the model relies solely on simple prior preferences (grey π) and connectivity between cortical and subcortical regions (grey C and θ), serving as controllers for inference and action selection. In contrast, the hierarchical structure adds specialized structures to the planning process (blue π , C and s), such as the observed functional role of frontal areas, enabling more efficient mapping between states, observations and mental actions. Learning is achieved by formulating a strategic plan involving precise policy selection based on anticipated action consequences. The hierarchical architecture effectively demonstrates learning by integrating precise modulation in policy selection, thereby enhancing confidence in cognitive actions and optimizing the decision-making process. In contrast, non-hierarchical structures exhibit unstable and variable action selection without precision, resulting in lower efficiency levels and indicating a lack of strategy.

the inclination to anticipate positive feedback and evade negative feedback. On the other hand, γ encapsulates the agent's prior beliefs concerning the level of confidence with which policies can be inferred, contingent upon the policy's free energy. Policy selection implies choosing the mental action (sequence of transition matrices) most likely to yield the expected outcomes and reduce the uncertainty of hidden states, a process influenced by prior preferences.

Given the novelty of active inference, there is limited literature on applications to NF. Mladenovic *et al.* [93] suggested that integrating active inference into a BCI exhibited significant improvements in performance compared to conventional methods. They observed variation between subjects and noted the potential for tuning the active inference's parameters to account for each subject's dynamics.

We constructed a self-regulation model based on active inference (figure 2b) to test scenarios about how top-down control could account for the variability in self-regulation learning, and the resolution of uncertainty when individuals seek to optimize their mental actions. Non-learners could be simulated by a model of a single layer guided by the prior probability of achieving an external goal, and connectivity between cortical and subcortical regions serving as controllers for inference and action selection (figure 2b, bottom). Landing on the brain state and mental action that explain the observed NF becomes an exhaustive computational task that is achieved by chance. Learners on the other hand are best modelled by a hierarchical structure connecting cortical areas, subcortical regions and frontal areas, allowing a more efficient access to the correct mental actions (figure 2b, top).

4. Neural substrates of brain self-regulation

One of the earliest studies that attempted to identify the neural correlates of self-regulation was conducted by Hinterberger *et al.* [94] using a slow cortical potential (SCP) BCI simultaneously with fMRI acquisition, showing that learning SCP control correlated with the activation of striatal and motor networks related to the associative binding of behaviour to the reward. Another study investigated the fMRI correlates of EEG-based brain self-regulation and identified the activation of the bilateral anterior insular cortex (AIC), anterior cingulate cortex (ACC), supplementary motor area (SMA), dorsomedial and lateral prefrontal cortex (DLPFC) and the posterior parietal cortex (PPC) [95]. A meta-analysis of fMRI NF studies showed similar results as in the above studies with correlated activity during self-regulation in AIC, ACC, SMA, DLPFC and PFC, PPC, basal ganglia and thalamus [96]. An interpretation of these results [5] attributed the different brain correlates of NF to a reward-processing network (comprising the ACC, AIC and ventral striatum), a control network (comprising the lateral occipital cortex (LOC), DLPFC, PPC and thalamus) and a learning network (involving the dorsal striatum). Animal studies have contributed substantially towards identifying the brain regions responsible for feedback learning, demonstrating, for example, that the complete blockade of NMDA receptors in the basal ganglia abolishes learning [59,97]. The brain's structural predictors of self-regulation and changes resulting from training have also been investigated (e.g. [98–100]).

A recent work by Vargas and colleagues conducted a direct evaluation of top-down and bottom-up processing of brain self-regulation, in the context of active inference [38] (figure 3). They performed dynamical causal modelling (DCM, [101]) of a real-time functional MRI NF dataset based on the work by Sepulveda *et al.* [37] to investigate the neural architecture that enables self-regulation learning and distinguishes learners from non-learners.

The authors evaluated the DCM models to assess which network structure is more likely to have generated the observed brain self-regulation data [37]. They also tested the difference in the brain networks of participants who successfully learned brain self-regulation and those who did not. According to their results, self-regulation learning was best modelled by a hierarchical architecture based on an active inference framework. In this architecture, the effects of feedback and self-regulation were mediated by top-down connections in the frontal cortex (including frontal superior and medial frontal gyri), anterior cingulate cortex and the striatum, together called the F-ACC-S axis. The results reveal different network modulation weights between NF learners and non-learners. These results could be viewed in the general context of learning, where active inference [90,101] suggests a shift from striatal bottom-up processing to frontal top-down influence as a possible explanation of learning [102].

It should be noted that active inference and RL models are not mutually exclusive. Active inference integrates a hierarchical neural architecture of RL-based learning by trial and error mediated by reward [103] with prior internal information used in generative predictions [102,104]. In this integrated framework, the generative model is updated by bottom-up sensory signals, thus increasing the predictive validity of the model that affects the top-down processes. This mechanism has been considered to result in better predictions and resolution of prediction errors originating in the striatum. The difference in top-down predictions based on frontal prior beliefs differentiates learners and non-learners. Based on this new view, relying only on bottom-up signals during self-regulation learning would be an untenable learning strategy due to the dimensionality of the

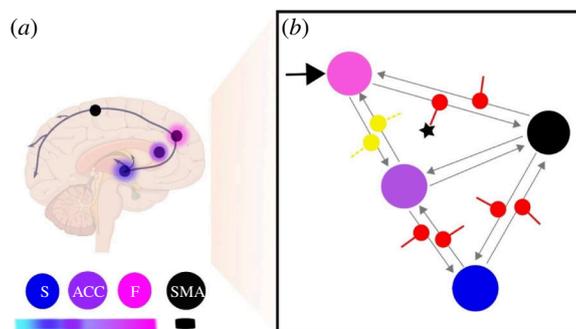


Figure 3. Self-regulation learning as active inference: dynamic causal modelling of an fMRI NF task (reproduced with permission from Vargas *et al.* [38]). A schematic illustration of the most probable model and parameter configuration of brain self-regulation learning. (a) The same network as (b) represented as brain regions, with nodes indicating the different brain regions and their hierarchical identification shown as colour labels following the functional anatomy of hierarchical active inference [90]. Different levels are represented by a colour gradient from fuchsia (frontal lobes, highest level) to blue-cyan (striatum). F, frontal lobe; ACC, anterior cingulate cortex; S, striatum; SMA, supplementary motor area. (b) The network connectivity during self-regulation and the significant difference between learners and non-learners. The input connection (C matrix) is indicated by a black arrow, while the effective connectivity architecture (Ae matrix) is represented by grey arrows. The modulation of extrinsic connectivity values (matrix Be) is depicted as red for excitatory and yellow for inhibitory. The star symbol denotes significant differences ($p < 0.05$) in connectivity between learners and non-learners. Specifically, it highlights variations in excitatory modulating connectivity between the frontal and target areas; with 'learners' corresponding to participants who effectively increase or decrease brain activity through NF training, evidenced by a significant increase in activity within the target region from the beginning to the end of the intervention, while non-learners do not exhibit this change.

search space and time constraints, and states that for self-regulation learning to succeed, top-down control may be necessary. This assertion is supported by findings in the cognitive functions that show that the internal generative model [90] is present in learning [105], abstract construction [106] and cognitive control [107].

However, both old and more recent work has shown neural activity in the context of brain self-regulation to be much more constrained than we might expect [31,75]. Simulations of fMRI NF have equally shown that, counterintuitively, the dimensionality of the search space may not be so high, leading to relatively fast learning [7,35,36]. Given that complex behaviours and information processing can be performed without awareness [108–111], there is substantial evidence suggesting we may not need conscious awareness, or explicit cognitive strategies, to exert control over our own neural activity. A probable scenario is that the brain has multiple (hierarchical) mechanisms for self-regulation. Depending on the NF experimental design, we may direct the brain into using one or the other mechanism.

5. Summary and conclusions

NF is a procedure that measures brain activity in real time and presents it as feedback to an individual, thus allowing them to self-regulate brain activity with effects on cognitive processes inferred from behaviour. One common argument is that NF studies can reveal how the measured brain activity causes a particular cognitive process. The causal claim often regards the measured brain activity being manipulated as an independent variable, analogous to brain stimulation. However, this causal inference is vulnerable to the argument that there are concurrent changes in other upstream brain activities that cause the changes in the activity generating the feedback. Alternately, NF may causally affect cognition by indirect means. We would argue that researchers should remain open to the idea that the trained brain activity could be part of a causal network that collectively affects cognition rather than being necessarily causally primary. That is, while the NF procedure might target one specific brain region, or one specific connection between two areas, changes in neural activity might ripple across an entire network, thus affecting other regions. The final outcome (cognitive or behavioural change) could depend on activity changes in areas downstream or separate from the NF target. Previous work with functional connectivity NF has indeed shown that post-NF, connectivity patterns beyond the target connection had also changed [112]. This inference may provide a better translation of evidence from NF studies to the rest of neuroscience. The recent advent of multivariate pattern analysis, combined with implicit NF, currently comprises the strongest case for causality [7,24,36,113]; however, future work will have to address this set of questions in depth.

RL and active inference are theoretical models to understand how the goal of optimizing behaviour and learning is achieved through experience. Active inference's prediction error aligns with RL's reward prediction error, driving updates to internal models. Active inference's exploratory drive mirrors RL's balance between exploration and exploitation. Both frameworks suggest that the brain learns by iteratively refining actions and predictions to achieve desired outcomes.

However, as a model for self-regulation via NF learning RL has fallen short of explaining why some people learn to self-regulate and others do not. RL focuses on maximizing rewards by adjusting actions based on positive or negative outcomes. In the context of self-regulation, this could be applied by considering regional brain activities as actions and the desired brain state as the goal. However, RL might not inherently account for the complex predictive processes and the balance between action, perception and prediction that are central to self-regulation, particularly in explicit NF paradigms. Active inference, on the other hand, explicitly incorporates the idea of minimizing prediction errors and optimizing internal models. It accounts for the interplay between predictions, actions and sensory input, making it well suited to explaining how individuals learn

to identify the correct internal actions for self-regulation. The rapid learning observed in self-regulation tasks like NF, where individuals need to identify the correct area for control amidst numerous potential states, can be attributed to a combination of cognitive mechanisms and neural processes, explicitly modelled by active inference in terms of minimization of prediction errors and optimizing internal models. Analyses based on large, pooled datasets across NF studies (e.g. [20,114]) will be critical in better adjudicating the computational processes underlying various NF approaches.

There are also deeper neuroscientific issues that are currently not well understood. Experimental findings in animal studies point to possible brain–behaviour decoupling because of NF training without simultaneous body–behaviour association [115]. Furthermore, the potential extinction of sensory–motor contingency has also been hypothesized with prolonged muscular inactivity in patients with complete paralysis, such as in the vegetative stage and other disorders of consciousness [99]. A phenomenon that may prevent longer term NF effects has been speculated to relate to the ‘rebound’ effect of neural activity resulting in a return to its original activation level after sustained NF training [5].

One limitation of the present approach to NF is related to the lack of bodily and environmental involvement in the NF paradigm. The lack of bodily and contextual settings and conditions that are conducive to brain–body–behaviour modulation during NF training could be the most important limitation of the present approaches to brain self-regulation.

Participants learn to control neural activity using mental strategies, such as emotional scenes, visual imagery, motor imagery and memory recall, without concomitant bodily movements, speech and overt physiological changes because of the restricted environment and the associated physiological artefacts of contemporary brain imaging technology. Furthermore, behavioural and physiological effects of NF are tested before and after training, with very little observation of ongoing changes in these outcomes. Studies are predominantly conducted in static, non-interactionist settings bereft of brain–body–environment interaction. Hence, real-world, socially and ecologically significant interactions are excluded from the contemporary NF approach. Future developments in NF should consider brain–body interactions inspired by embodied–enactive cognition theories. In the proposed approach, the participant is trained in the natural environmental context, which operates in daily life and constrains brain function and behaviour. Embodied cognition (e.g. [116–118]) can be succinctly described as the cognitive processing that includes and extends beyond the agent’s brain in the agent’s ongoing dynamic interactions with the environment. This view begs for innovative techniques to integrate portable and wearable neurophysiological monitoring technology for real-time feedback to train brain and body to act together in a contextually meaningful manner.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. This article has no additional data.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors’ contributions. R.S.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, visualization, writing—original draft, writing—review and editing; A.S.-C.: conceptualization, investigation, resources, writing—original draft, writing—review and editing; G.V.: investigation, visualization, writing—original draft, writing—review and editing; A.C.: writing—original draft, writing—review and editing; W.E.-D.: writing—original draft, writing—review and editing; A.J.: conceptualization, investigation, writing—original draft, writing—review and editing; E.F.: conceptualization, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. RS was supported by intramural funding from St. Jude Children’s Research Hospital. WeD acknowledges the support of ANID, Chile (projects FONDECYT 1241695, ANILLO ACT210053, Basal FB0008), ValgrAI, and the Generalitat Valenciana, Spain. A.C. was partially supported by a KDDI collaborative research contract; by the Innovative Science and Technology Initiative for Security (grant number JPJ004596) from ATLA, Japan; by the Ikegaya Brain-AI Hybrid ERATO (grant number JPMJER1801) from the Japan Science and Technology Agency; and by JSPS KAKENHI (grant number JP22H05156). E.F. was supported by NIH grant NS12542.

References

1. Chamola V, Vineet A, Nayyar A, Hossain E. 2020 Brain–computer interface-based humanoid control: a review. *Sensors (Basel)* **20**, 3620. (doi:10.3390/s20133620)
2. Choi JR, Kim SM, Ryu RH, Kim SP, Sohn JW. 2018 Implantable neural probes for brain–machine interfaces—current developments and future prospects. *Exp. Neurol.* **27**, 453–471. (doi:10.5607/en.2018.27.6.453)
3. Lebedev MA, Nicolelis MAL. 2017 Brain–machine interfaces: from basic science to neuroprostheses and neurorehabilitation. *Physiol. Rev.* **97**, 767–837. (doi:10.1152/physrev.00027.2016)
4. Mridha MF, Das SC, Kabir MM, Lima AA, Islam MR, Watanobe Y. 2021 Brain–computer interface: advancement and challenges. *Sensors (Basel)* **21**, 5746. (doi:10.3390/s21175746)
5. Sitaram R *et al.* 2017 Closed-loop brain training: the science of neurofeedback. *Nat. Rev. Neurosci.* **18**, 86–100. (doi:10.1038/nrn.2016.164)
6. Sulzer J *et al.* 2013 Real-time fMRI neurofeedback: progress and challenges. *Neuroimage* **76**, 386–399. (doi:10.1016/j.neuroimage.2013.03.033)
7. Watanabe T, Sasaki Y, Shibata K, Kawato M. 2017 Advances in fMRI real-time neurofeedback. *Trends Cogn. Sci.* **21**, 997–1010. (doi:10.1016/j.tics.2017.09.010)
8. Loriette C, Ziane C, Ben Hamed S. 2021 Neurofeedback for cognitive enhancement and intervention and brain plasticity. *Rev. Neurol. (Paris)* **177**, 1133–1144. (doi:10.1016/j.neurol.2021.08.004)
9. Marzbani H, Marateb HR, Mansourian M. 2016 Neurofeedback: a comprehensive review on system design. *Basic Clin. Neurosci.* **7**, 143–158. (doi:10.15412/J.BCN.03070208)
10. Pindi P, Houenou J, Piguet C, Favre P. 2022 Real-time fMRI neurofeedback as a new treatment for psychiatric disorders: a meta-analysis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **119**, 110605. (doi:10.1016/j.pnpbp.2022.110605)
11. Fede SJ, Dean SF, Manuweera T, Momenan R. 2020 A guide to literature informed decisions in the design of real time fMRI neurofeedback studies: a systematic review. *Front. Hum. Neurosci.* **14**, 60. (doi:10.3389/fnhum.2020.00060)
12. Thibault RT, MacPherson A, Lifshitz M, Roth RR, Raz A. 2018 Neurofeedback with fMRI: a critical systematic review. *Neuroimage* **172**, 786–807. (doi:10.1016/j.neuroimage.2017.12.071)

13. Haug A *et al.* 2021 Predictors of real-time fMRI neurofeedback performance and improvement—a machine learning mega-analysis. *Neuroimage* **237**, 118207. (doi:10.1016/j.neuroimage.2021.118207)
14. Price J, Budzynski T. 2009 Anxiety, EEG patterns, and neurofeedback. In *Introduction to quantitative EEG and neurofeedback* (eds TH Budzynski, HK Budzynski, JR Evans, A Abarbanel), pp. 453–472. New York, NY: Academic Press. (doi:10.1016/B978-0-12-374534-7.00017-4)
15. Davelaar EJ. 2018 Mechanisms of neurofeedback: a computation-theoretic approach. *Neuroscience* **378**, 175–188. (doi:10.1016/j.neuroscience.2017.05.052)
16. Epstein LH, Blanchard EB. 1977 Biofeedback, self-control, and self-management. *Biofeedback Self Regul.* **2**, 201–211. (doi:10.1007/BF00998670)
17. Gaume A, Vialatte A, Mora-Sánchez A, Ramdani C, Vialatte FB. 2016 A psychoengineering paradigm for the neurocognitive mechanisms of biofeedback and neurofeedback. *Neurosci. Biobehav. Rev.* **68**, 891–910. (doi:10.1016/j.neubiorev.2016.06.012)
18. Leigh H. 1978 Self-control, biofeedback, and change in 'psychosomatic' approach. *Psychother. Psychosom.* **30**, 130–136. (doi:10.1159/000287291)
19. Renton AI, Painter DR, Mattingley JB. 2021 Implicit neurofeedback training of feature-based attention promotes biased sensory processing during integrative decision-making. *J. Neurosci.* **41**, 8233–8248. (doi:10.1523/JNEUROSCI.0243-21.2021)
20. Cortese A *et al.* 2021 The DecNef collection, fMRI data from closed-loop decoded neurofeedback experiments. *Sci. Data* **8**, 65. (doi:10.1038/s41597-021-00845-7)
21. Krause F, Benjamins C, Lührs M, Eck J, Noirhomme Q, Rosenke M, Brunheim S, Sorger B, Goebel R. 2017 Real-time fMRI-based self-regulation of brain activation across different visual feedback presentations. *Brain-Comput. Interfaces* **4**, 87–101. (doi:10.1080/2326263X.2017.1307096)
22. Nacke LE. 2010 Wiimote vs. controller: electroencephalographic measurement of affective gameplay interaction. In *Proc. Int. Acad. Conf. on the Future of Game Design and Technology*, pp. 159–166. Vancouver British Columbia Canada: ACM. (doi:10.1145/1920778.1920801)
23. Ramot M, Grossman S, Friedman D, Malach R. 2016 Covert neurofeedback without awareness shapes cortical network spontaneous connectivity. *Proc. Natl Acad. Sci. USA* **113**, E2413–E2420. (doi:10.1073/pnas.1516857113)
24. Muñoz-Moldes S, Cleeremans A. 2020 Delineating implicit and explicit processes in neurofeedback learning. *Neurosci. Biobehav. Rev.* **118**, 681–688. (doi:10.1016/j.neubiorev.2020.09.003)
25. Fetz EE, Baker MA. 1973 Operantly conditioned patterns on precentral unit activity and correlated responses in adjacent cells and contralateral muscles. *J. Neurophysiol.* **36**, 179–204. (doi:10.1152/jn.1973.36.2.179)
26. Fetz EE, Finocchio DV. 1971 Operant conditioning of specific patterns of neural and muscular activity. *Science* **174**, 431–435. (doi:10.1126/science.174.4007.431)
27. Fetz EE, Wyler AR. 1973 Operantly conditioned firing patterns of epileptic neurons in the monkey motor cortex. *Exp. Neurol.* **40**, 586–607. (doi:10.1016/0014-4886(73)90098-8)
28. Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, Patil PG, Henriquez CS, Nicolelis MAL. 2003 Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol.* **1**, E42. (doi:10.1371/journal.pbio.0000042)
29. Jarosiewicz B, Chase SM, Fraser GW, Velliste M, Kass RE, Schwartz AB. 2008 Functional network reorganization during learning in a brain-computer interface paradigm. *Proc. Natl Acad. Sci. USA* **105**, 19486–19491. (doi:10.1073/pnas.0808113105)
30. Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, Yu BM, Batista AP. 2014 Neural constraints on learning. *Nature* **512**, 423–426. (doi:10.1038/nature13665)
31. Golub MD, Sadtler PT, Oby ER, Quick KM, Ryu SI, Tyler-Kabara EC, Batista AP, Chase SM, Yu BM. 2018 Learning by neural reassociation. *Nat. Neurosci.* **21**, 607–616. (doi:10.1038/s41593-018-0095-3)
32. Oby ER, Golub MD, Hennig JA, Degenhart AD, Tyler-Kabara EC, Yu BM, Chase SM, Batista AP. 2019 New neural activity patterns emerge with long-term learning. *Proc. Natl Acad. Sci. USA* **116**, 15210–15215. (doi:10.1073/pnas.1820296116)
33. Athalye VR, Ganguly K, Costa RM, Carmena JM. 2017 Emergence of coordinated neural dynamics underlies neuroprosthetic learning and skillful control. *Neuron* **93**, 955–970. (doi:10.1016/j.neuron.2017.01.016)
34. Bagdasaryan J, Quyen MLV. 2013 Experiencing your brain: neurofeedback as a new bridge between neuroscience and phenomenology. *Front. Hum. Neurosci.* **7**, 680. (doi:10.3389/fnhum.2013.00680)
35. Oblak EF, Lewis-Peacock JA, Sulzer JS. 2017 Self-regulation strategy, feedback timing and hemodynamic properties modulate learning in a simulated fMRI neurofeedback environment. *PLoS Comput. Biol.* **13**, e1005681. (doi:10.1371/journal.pcbi.1005681)
36. Shibata K, Lisi G, Cortese A, Watanabe T, Sasaki Y, Kawato M. 2019 Toward a comprehensive understanding of the neural mechanisms of decoded neurofeedback. *Neuroimage* **188**, 539–556. (doi:10.1016/j.neuroimage.2018.12.022)
37. Sepulveda P, Sitaram R, Rana M, Montalba C, Tejos C, Ruiz S. 2016 How feedback, motor imagery, and reward influence brain self-regulation using real-time fMRI. *Hum. Brain Mapp.* **37**, 3153–3171. (doi:10.1002/hbm.23228)
38. Vargas G, Araya D, Sepulveda P, Rodriguez-Fernandez M, Friston KJ, Sitaram R, El-Derey W. 2023 Self-regulation learning as active inference: dynamic causal modeling of an fMRI neurofeedback task. *Front. Neurosci.* **17**, 1212549. (doi:10.3389/fnins.2023.1212549)
39. Lubianiker N, Paret C, Dayan P, Hendler T. 2022 Neurofeedback through the lens of reinforcement learning. *Trends Neurosci.* **45**, 579–593. (doi:10.1016/j.tins.2022.03.008)
40. Egner T, Sterman MB. 2006 Neurofeedback treatment of epilepsy: from basic rationale to practical application. *Expert Rev. Neurother.* **6**, 247–257. (doi:10.1586/14737175.6.2.247)
41. Cortese A, Amano K, Koizumi A, Lau H, Kawato M. 2017 Decoded fMRI neurofeedback can induce bidirectional confidence changes within single participants. *Neuroimage* **149**, 323–337. (doi:10.1016/j.neuroimage.2017.01.069)
42. Wood G, Kober SE, Witte M, Neuper C. 2014 On the need to better specify the concept of 'control' in brain-computer-interfaces/neurofeedback research. *Front. Syst. Neurosci.* **8**, 171. (doi:10.3389/fnsys.2014.00171)
43. Alkoby O, Abu-Rmieleh A, Shriki O, Todder D. 2018 Can we predict who will respond to neurofeedback? A review of the inefficacy problem and existing predictors for successful EEG neurofeedback learning. *Neuroscience* **378**, 155–164. (doi:10.1016/j.neuroscience.2016.12.050)
44. Johnson KA, Hartwell K, LeMatty T, Borckardt J, Morgan PS, Govindarajan K, Brady K, George MS. 2012 Intermittent 'real-time' fMRI feedback is superior to continuous presentation for a motor imagery task: a pilot study. *J. Neuroimaging* **22**, 58–66. (doi:10.1111/j.1552-6569.2010.00529.x)
45. Skottnik L, Sorger B, Kamp T, Linden D, Goebel R. 2019 Success and failure of controlling the real-time functional magnetic resonance imaging neurofeedback signal are reflected in the striatum. *Brain Behav.* **9**, e01240. (doi:10.1002/brb3.1240)
46. Gevensleben H, Moll GH, Rothenberger A, Heinrich H. 2014 Neurofeedback in attention-deficit/hyperactivity disorder—different models, different ways of application. *Front. Hum. Neurosci.* **8**, 846. (doi:10.3389/fnhum.2014.00846)
47. Omejc N, Rojc B, Battaglini PP, Marusic U. 2018 Review of the therapeutic neurofeedback method using electroencephalography: EEG neurofeedback. *Bosn. J. Basic Med. Sci.* (doi:10.17305/bjbm.2018.3785)
48. Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. 2001 Neurophysiological investigation of the basis of the fMRI signal. *Nat. New Biol.* **412**, 150–157. (doi:10.1038/35084005)

49. Heeger DJ, Ress D. 2002 What does fMRI tell us about neuronal activity? *Nat. Rev. Neurosci.* **3**, 142–151. (doi:10.1038/nrn730)
50. Handwerker DA, Ollinger JM, D'Esposito M. 2004 Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* **21**, 1639–1651. (doi:10.1016/j.neuroimage.2003.11.029)
51. Naseer N, Hong KS. 2015 fNIRS-based brain-computer interfaces: a review. *Front. Hum. Neurosci.* **9**. (doi:10.3389/fnhum.2015.00003)
52. Uchitel J, Vidal-Rosas EE, Cooper RJ, Zhao H. 2021 Wearable, integrated EEG-fNIRS technologies: a review. *Sensors (Basel)* **21**, 6106. (doi:10.3390/s21186106)
53. Buxton RB, Uludağ K, Dubowitz DJ, Liu TT. 2004 Modeling the hemodynamic response to brain activation. *Neuroimage* **23**, S220–S233. (doi:10.1016/j.neuroimage.2004.07.013)
54. Polimeni JR, Lewis LD. 2021 Imaging faster neural dynamics with fast fMRI: a need for updated models of the hemodynamic response. *Prog. Neurobiol.* **207**, 102174. (doi:10.1016/j.pneurobio.2021.102174)
55. Caligiore D, Arbib MA, Miall RC, Baldassarre G. 2019 The super-learning hypothesis: integrating learning processes across cortex, cerebellum and basal ganglia. *Neurosci. Biobehav. Rev.* **100**, 19–34. (doi:10.1016/j.neubiorev.2019.02.008)
56. Galea JM, Vazquez A, Pasricha N, de Xivry JJO, Celnik P. 2011 Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. *Cereb. Cortex* **21**, 1761–1770. (doi:10.1093/cercor/bhq246)
57. Bao S, Lei Y. 2022 Memory decay and generalization following distinct motor learning mechanisms. *J. Neurophysiol.* **128**, 1534–1545. (doi:10.1152/jn.00105.2022)
58. Uehara S, Mawase F, Celnik P. 2018 Learning similar actions by reinforcement or sensory-prediction errors rely on distinct physiological mechanisms. *Cereb. Cortex* **28**, 3478–3490. (doi:10.1093/cercor/bhx214)
59. Koralek AC, Jin X, Long JD, Costa RM, Carmena JM. 2012 Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. *Nat. New Biol.* **483**, 331–335. (doi:10.1038/nature10845)
60. Neely RM, Koralek AC, Athalye VR, Costa RM, Carmena JM. 2018 Volitional modulation of primary visual cortex activity requires the basal ganglia. *Neuron* **97**, 1356–1368. (doi:10.1016/j.neuron.2018.01.051)
61. Eaton RW, Libey T, Fetz EE. 2017 Operant conditioning of neural activity in freely behaving monkeys with intracranial reinforcement. *J. Neurophysiol.* **117**, 1112–1125. (doi:10.1152/jn.00423.2016)
62. Schönauer M, Geisler T, Gais S. 2014 Strengthening procedural memories by reactivation in sleep. *J. Cogn. Neurosci.* **26**, 143–153. (doi:10.1162/jocn_a_00471)
63. Gulati T, Ramanathan DS, Wong CC, Ganguly K. 2014 Reactivation of emergent task-related ensembles during slow-wave sleep after neuroprosthetic learning. *Nat. Neurosci.* **17**, 1107–1113. (doi:10.1038/nn.3759)
64. Black AH, Cott A, Pavloski R. 1977 The operant learning theory approach to biofeedback training. In *Biofeedback. Theory and research* (eds G Schwartz, J Beatty), pp. 89–127. New York, NY: Academic Press.
65. Ros T *et al.* 2020 Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist). *Brain* **143**, 1674–1685. (doi:10.1093/brain/awaa009)
66. Black AH, Cott A. 1977 Biofeedback and behavior. In *A perspective on biofeedback* (eds J Beatty, H Legewie), pp. 7–19. Boston, MA: Springer US. (doi:10.1007/978-1-4684-2526-0_2)
67. Lepper MR, Greene D, Nisbett RE. 1973 Undermining children's intrinsic interest with extrinsic reward: a test of the 'overjustification' hypothesis. *J. Pers. Soc. Psychol.* **28**, 129–137. (doi:10.1037/h0035519)
68. Kadosh KC, Staunton G. 2019 A systematic review of the psychological factors that influence neurofeedback learning outcomes. *Neuroimage* **185**, 545–555. (doi:10.1016/j.neuroimage.2018.10.021)
69. Balleine BW, Dickinson A. 1998 Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407–419. (doi:10.1016/s0028-3908(98)00033-1)
70. Dayan P, Berridge KC. 2014 Model-based and model-free Pavlovian reward learning: reevaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 473–492. (doi:10.3758/s13415-014-0277-8)
71. Daw ND, Niv Y, Dayan P. 2005 Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711. (doi:10.1038/nn1560)
72. Doya K, Samejima K, Katagiri K-i, Kawato M. 2002 Multiple model-based reinforcement learning. *Neural Comput.* **14**, 1347–1369. (doi:10.1162/089976602753712972)
73. Montague PR, Dayan P, Sejnowski TJ. 1996 A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947. (doi:10.1523/JNEUROSCI.16-05-01936.1996)
74. Sutton RS, Barto AG. 1990 Time-derivative models of Pavlovian reinforcement. In *Learning and computational neuroscience: foundations of adaptive networks* (eds M Gabriel, J Moore), pp. 497–537. Cambridge, MA: The MIT Press.
75. Luczak A, Barthó P, Harris KD. 2009 Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* **62**, 413–425. (doi:10.1016/j.neuron.2009.03.014)
76. Oblak EF, Sulzer JS, Lewis-Peacock JA. 2019 A simulation-based approach to improve decoded neurofeedback performance. *Neuroimage* **195**, 300–310. (doi:10.1016/j.neuroimage.2019.03.062)
77. Chase HW, Kumar P, Eickhoff SB, Dombrowski AY. 2015 Reinforcement learning models and their neural correlates: an activation likelihood estimation meta-analysis. *Cogn. Affect. Behav. Neurosci.* **15**, 435–459. (doi:10.3758/s13415-015-0338-7)
78. Schultz W, Dayan P, Montague PR. 1997 A neural substrate of prediction and reward. *Science* **275**, 1593–1599. (doi:10.1126/science.275.5306.1593)
79. Doya K. 1999 What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw.* **12**, 961–974. (doi:10.1016/s0893-6080(99)00046-5)
80. Paret C, Zaehring J, Ruf M, Ende G, Schmahl C. 2019 The orbitofrontal cortex processes neurofeedback failure signals. *Behav. Brain Res.* **369**, 111938. (doi:10.1016/j.bbr.2019.111938)
81. Zhao Z, Yao S, Zwerings J, Zhou X, Zhou F, Kendrick KM, Chen H, Mathiak K, Becker B. 2021 Putamen volume predicts real-time fMRI neurofeedback learning success across paradigms and neurofeedback target regions. *Hum. Brain Mapp.* **42**, 1879–1887. (doi:10.1002/hbm.25336)
82. Athalye VR, Carmena JM, Costa RM. 2020 Neural reinforcement: re-entering and refining neural dynamics leading to desirable outcomes. *Curr. Opin. Neurobiol.* **60**, 145–154. (doi:10.1016/j.conb.2019.11.023)
83. Law CT, Gold JL. 2009 Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat. Neurosci.* **12**, 655–663. (doi:10.1038/nn.2304)
84. Moran RJ, Symmonds M, Dolan RJ, Friston KJ. 2014 The brain ages optimally to model its environment: evidence from sensory learning over the adult lifespan. *PLoS Comput. Biol.* **10**, e1003422. (doi:10.1371/journal.pcbi.1003422)

85. Friston K, FitzGerald T, Rigoli F, Schwartenbeck P, O'Doherty J, Pezzulo G. 2016 Active inference and learning. *Neurosci. Biobehav. Rev.* **68**, 862–879. (doi:10.1016/j.neubiorev.2016.06.022)
86. Friston K, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G. 2017 Active inference: a process theory. *Neural Comput.* **29**, 1–49. (doi:10.1162/NECO_a_00912)
87. Pezzulo G, Rigoli F, Friston KJ. 2018 Hierarchical active inference: a theory of motivated control. *Trends Cogn. Sci.* **22**, 294–306. (doi:10.1016/j.tics.2018.01.009)
88. Chen AG, Benrimoh D, Parr T, Friston KJ. 2020 A Bayesian account of generalist and specialist formation under the active inference framework. *Front. Artif. Intell.* **3**, 69. (doi:10.3389/frai.2020.00069)
89. Smith R, Friston KJ, Whyte CJ. 2022 A step-by-step tutorial on active inference and its application to empirical data. *J. Math. Psychol.* **107**, 102632. (doi:10.1016/j.jmp.2021.102632)
90. Pezzulo G, Rigoli F, Friston K. 2015 Active inference, homeostatic regulation and adaptive behavioural control. *Prog. Neurobiol.* **134**, 17–35. (doi:10.1016/j.pneurobio.2015.09.001)
91. Da Costa L, Parr T, Sajid N, Veselic S, Neacsu V, Friston K. 2020 Active inference on discrete state-spaces: a synthesis. *J. Math. Psychol.* **99**, 102447. (doi:10.1016/j.jmp.2020.102447)
92. Friston KJ, Rosch R, Parr T, Price C, Bowman H. 2017 Deep temporal models and active inference. *Neurosci. Biobehav. Rev.* **77**, 388–402. (doi:10.1016/j.neubiorev.2017.04.009)
93. Mladenovic J, Frey J, Joffily M, Maby E, Lotte F, Mattout J. 2020 Active inference as a unifying, generic and adaptive framework for a P300-based BCI. *J. Neural Eng.* **17**, 016054. (doi:10.1088/1741-2552/ab5d5c)
94. Hinterberger T, Veit R, Wilhelm B, Weiskopf N, Vatine JJ, Birbaumer N. 2005 Neuronal mechanisms underlying control of a brain–computer interface. *Eur. J. Neurosci.* **21**, 3169–3181. (doi:10.1111/j.1460-9568.2005.04092.x)
95. Ninaus M, Kober SE, Witte M, Koschutnig K, Stangl M, Neuper C, Wood G. 2013 Neural substrates of cognitive control under the belief of getting neurofeedback training. *Front. Hum. Neurosci.* **7**, 914. (doi:10.3389/fnhum.2013.00914)
96. Emmert K *et al.* 2016 Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: how is brain regulation mediated? *Neuroimage* **124**, 806–812. (doi:10.1016/j.neuroimage.2015.09.042)
97. Koralek AC, Costa RM, Carmena JM. 2013 Temporally precise cell-specific coherence develops in corticostriatal networks during learning. *Neuron* **79**, 865–872. (doi:10.1016/j.neuron.2013.06.047)
98. Scholz J, Klein MC, Behrens TEJ, Johansen-Berg H. 2009 Training induces changes in white-matter architecture. *Nat. Neurosci.* **12**, 1370–1371. (doi:10.1038/nn.2412)
99. 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. (doi:10.1177/1550059413476031)
100. Kasahara K, DaSalla CS, Honda M, Hanakawa T. 2015 Neuroanatomical correlates of brain–computer interface performance. *Neuroimage* **110**, 95–100. (doi:10.1016/j.neuroimage.2015.01.055)
101. Friston K, Moran R, Seth AK. 2013 Analysing connectivity with Granger causality and dynamic causal modelling. *Curr. Opin. Neurobiol.* **23**, 172–178. (doi:10.1016/j.conb.2012.11.010)
102. Smith R, Schwartenbeck P, Parr T, Friston KJ. 2020 An active inference approach to modeling structure learning: concept learning as an example case. *Front. Comput. Neurosci.* **14**, 41. (doi:10.3389/fncom.2020.00041)
103. Mnih V *et al.* 2015 Human-level control through deep reinforcement learning. *Nature* **518**, 529–533. (doi:10.1038/nature14236)
104. Parr T, Corcoran AW, Friston KJ, Hohwy J. 2019 Perceptual awareness and active inference. *Neurosci. Conscious.* **2019**, niz012. (doi:10.1093/nc/niz012)
105. Huang Y, Yaple ZA, Yu R. 2020 Goal-oriented and habitual decisions: neural signatures of model-based and model-free learning. *Neuroimage* **215**, 116834. (doi:10.1016/j.neuroimage.2020.116834)
106. Collins AGE, Frank MJ. 2013 Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol. Rev.* **120**, 190–229. (doi:10.1037/a0030852)
107. Bassett DS, Yang M, Wymbs NF, Grafton ST. 2015 Learning-induced autonomy of sensorimotor systems. *Nat. Neurosci.* **18**, 744–751. (doi:10.1038/nn.3993)
108. Goldberg RF, Perfetti CA, Schneider W. 2006 Perceptual knowledge retrieval activates sensory brain regions. *J. Neurosci.* **26**, 4917–4921. (doi:10.1523/JNEUROSCI.5389-05.2006)
109. van Gaal S, Ridderinkhof KR, Scholte HS, Lamme VAF. 2010 Unconscious activation of the prefrontal no-go network. *J. Neurosci.* **30**, 4143–4150. (doi:10.1523/JNEUROSCI.2992-09.2010)
110. Charles L, Gaillard R, Amado I, Krebs MO, Bendjema N, Dehaene S. 2017 Conscious and unconscious performance monitoring: evidence from patients with schizophrenia. *Neuroimage* **144**, 153–163. (doi:10.1016/j.neuroimage.2016.09.056)
111. Cortese A, Lau H, Kawato M. 2020 Unconscious reinforcement learning of hidden brain states supported by confidence. *Nat. Commun.* **11**, 4429. (doi:10.1038/s41467-020-17828-8)
112. Megumi F, Yamashita A, Kawato M, Imamizu H. 2015 Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Front. Hum. Neurosci.* **9**, 160. (doi:10.3389/fnhum.2015.00160)
113. Taschereau-Dumouchel V, Kawato M, Lau H. 2020 Multivoxel pattern analysis reveals dissociations between subjective fear and its physiological correlates. *Mol. Psychiatry* **25**, 2342–2354. (doi:10.1038/s41380-019-0520-3)
114. Haug A *et al.* 2020 Can we predict real-time fMRI neurofeedback learning success from pretraining brain activity? *Hum. Brain Mapp.* **41**, 3839–3854. (doi:10.1002/hbm.25089)
115. Fetz EE, Perlmutter SI, Prut Y, Seki K, Votaw S. 2002 Roles of primate spinal interneurons in preparation and execution of voluntary hand movement. *Brain Res. Rev.* **40**, 53–65. (doi:10.1016/s0165-0173(02)00188-1)
116. Farina M. 2021 Embodied cognition: dimensions, domains and applications. *Adapt. Behav.* **29**, 73–88. (doi:10.1177/1059712320912963)
117. Foglia L, Wilson RA. 2013 Embodied cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* **4**, 319–325. (doi:10.1002/wcs.1226)
118. Goldman A, de Vignemont F. 2009 Is social cognition embodied? *Trends Cogn. Sci.* **13**, 154–159. (doi:10.1016/j.tics.2009.01.007)