Modeling subject perception and behaviour during neurofeedback training

Côme Annicchiarico, Fabien Lotte, Jérémie Mattout

Modeling subject perception and behaviour during neurofeedback training

Côme ANNICCHIARICO[1] , Fabien LOTTE[2] , Jérémie MATTOUT[1,3]

1) Lyon Neuroscience Research Center, CRNL, INSERM U1028, CNRS UMR5292, Computation, Cognition and Neurophysiology Team, Lyon, France
2) Inria Bordeaux Sud-Ouest / LaBRi (Univ; Bordeaux – CNRS – Bordeaux INP)
3) Université Claude Bernard Lyon 1, Lyon, France

Keywords: Neurofeedback, Modeling, Active Inference

Abstract

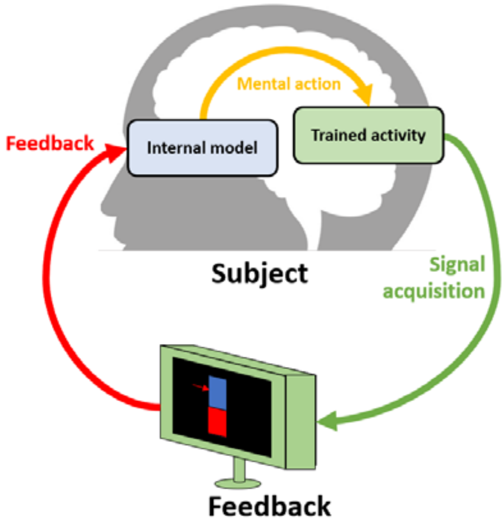
Neurofeedback training (NFT) describes a closed-loop paradigm in which a subject is provided with a real time evaluation of his/her brain activity. As a learning process, it is designed to help the subject learn to apprehend his/her own cognitive states and better modulate them through mental actions. Its use for therapeutic purposes has gained a lot of traction in the public sphere in the last decade, but conflicting evidence concerning its efficacy has led to a two-pronged effort from the scientific community. First, a call for experimental protocols and reports standardization [1], aiming to reduce the variability of the results and provide a reliable set of data to describe empirical findings. Second, an effort towards a formal description of the neurofeedback loop and the main hypotheses that guide the design of our experiments, in order to explain or even predict the effects of such training [2,3].

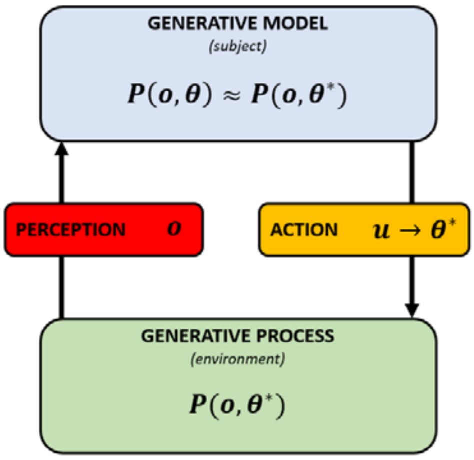
This work intends to contribute to the second effort by proposing a mathematical formalization of the mechanisms at play in this complex dyadic dynamical system.

A typical Neurofeedback experiment aims at having a significant (positive) impact on behavioral symptoms (e.g. focusing and learning abilities in children with attention disorder). Therefore, strong hypotheses are made about the relationship between a related mental state (e.g. the ongoing covert attentional effort) and a neurophysiological marker (e.g. the ratio between theta and beta band power as measured with EEG [4]). In addition to such hypotheses made by the experimenter, it is also important to account for the subject's beliefs (e.g., his trust in the feedback), which partly depend upon the provided instructions, and will impact her/his expectations and the dynamics of training.

Our formulation makes those hypotheses explicit, in a quantitative manner, so that one can simulate the ensuing closed-loop interaction (Fig. 1.a) whose aim is the self-regulation of the targeted neuromarker. As such, it enables us to evaluate the consequences of various approximate or even erroneous hypotheses (e.g. targeting an inappropriate physiological marker [4], providing a suboptimal feedback,…).

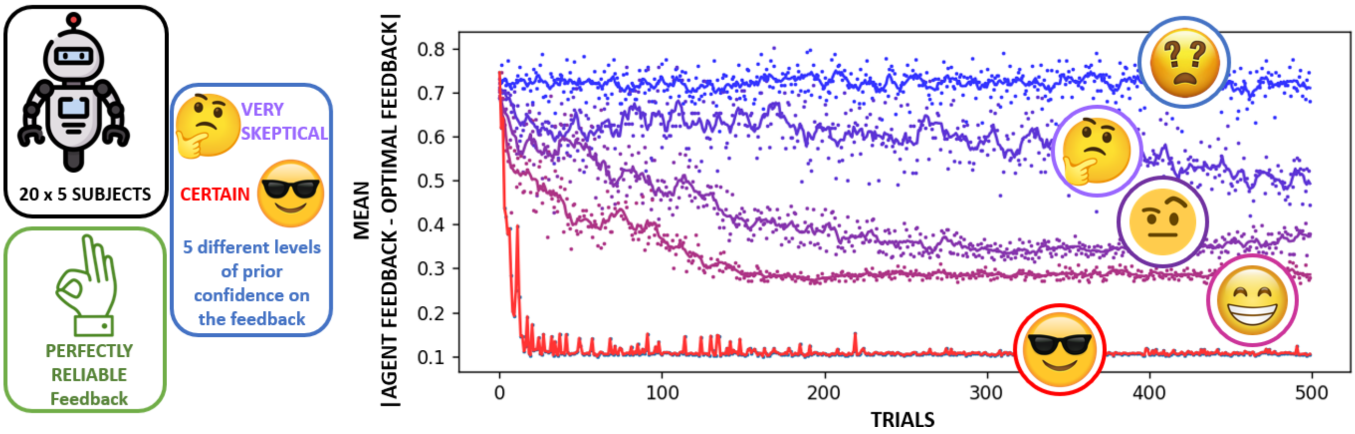
Our model relies on the Bayesian framework. It enables to instantiate an agent who entertains and updates a probabilistic (generative) model of the Neurofeedback environment. This rests on the assumption that the agent can only infer its own mental state through sensory feedback and the knowledge of its own mental actions aiming at modifying that state. Hence, using Bayesian computation, we cast the Neurofeedback training as the active process, for a subject, of inferring her/his mental states as well as the environment dynamics despite multiple significant sources of uncertainty (Fig 1.b). To simulate this loop and the effect of various sources of uncertainty, we take





a. Neurofeedback paradigms measure subject physiological activity (green) and process it to design a feedback (red). The subject tries to make sense of the feedback and learn (blue) how to act upon it (yellow) to achieve an objective.

b. We cast the neurofeedback subject experience as a high uncertainty inference problem where he/she tries to figure out the hidden states of the world (hidden mental states θ^*) that caused the feedback observed (o) and the optimal actions to get better results (u). The subject gets better at it by learning the dynamics of the environment (Perception + action).



c. We simulate the feedback performance of various artificial agents across trials performing neurofeedback training, starting with different prior confidence levels regarding feedback reliability. We show that even if the feedback is truly reliable, agents with excessive initial skepticism (blue) perform much poorer than those with high/absolute confidence in the feedback. (red)

Figure 1. Our approach formalizes the various functional components of the neurofeedback loop (a.), relies on Bayesian formalism to cast it as an inference problem (b.) and uses Active Inference to simulate the performance of artificial agents with varying initial parameters (c.).

advantage of the Active Inference framework [5,6], a Bayesian approach to belief updating that provides a biologically plausible model of perception, action and learning. The framework proposes an account of how independent systems balance out explorative and exploitative behaviour to achieve their goals, which is crucial with NFT paradigms.

We introduce a generic NFT task and we simulate the evolution of Active-Inference embedded artificial agents performing this training with various initial parameters (initial beliefs, motivation, habits, feedback quality, etc.). Agents are tasked with learning the reliability of the feedback but also the effect of their own actions on the mental states. In a first illustrative example, we show that training efficacy drops quickly with feedback quality, but we also find that a perfect feedback signal is not quite enough to guarantee training success, even in a very simplified representation. By changing initial agent confidence about the feedback as well as prior knowledge about the effect of its actions, we can emulate very different learning trajectories. Interestingly, we found that although the feedback may be perfect, excessive skepticism about it yields poor performance when facing too much uncertainty in the action model (Fig 1.c). Ongoing work consists in exploring the effects of various endogenous (e.g. motivation) and exogenous (e.g. feedback design) factors onto mental training.

References

[1] Ros et al., Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist). Brain, 143(6):1674–1685, June 2020.

[2] Oblak et al., Self-regulation strategy, feedback timing and hemodynamic properties modulate learning in a simulated fMRI neurofeedback environment. PLoS computational biology, 13:e1005681, July 2017.

[3] Davelaar, Mechanisms of Neurofeedback: A Computation-theoretic Approach. Neuroscience, 378:175–188, May 2018.

[4] Arns, M., Clark, C.R., Trullinger, M. et al. Neurofeedback and Attention-Deficit/Hyperactivity-Disorder (ADHD) in Children: Rating the Evidence and Proposed Guidelines. Appl Psychophysiol Biofeedback 45, 39–48 (2020). https://doi.org/10.1007/s10484-020-09455-2

[5] Friston et al., Active Inference: A Process Theory. Neural Computation, 29(1):1–49, January 2017.

[6] Friston et al., Active inference and learning. Neuroscience & Biobehavioral Reviews, 68:862 – 879,September 2016.