Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Transcranial direct current stimulation over the left prefrontal cortex increases randomness of choice in instrumental learning

Zsolt Turi [a,*,1], Matthias Mittner [b,1], Alexander Opitz [a,1], Miriam Popkes [a], Walter Paulus [a] and Andrea Antal [a]

[a] Department Clinical Neurophysiology, University Medical Center, Georg-August University, Göttingen, Germany
[b] Department of Psychology, University of Tromsø, Norway

ARTICLE INFO

Article history:
Received 6 January 2014
Reviewed 1 April 2014
Revised 13 May 2014
Accepted 26 August 2014
Action editor Jacinta O'Shea
Published online 11 September 2014

Keywords:
Transcranial direct current stimulation
Dorsolateral prefrontal cortex
Probabilistic learning task
Exploration
Exploitation
Working memory

ABSTRACT

Introduction: There is growing evidence from neuro-computational studies that instrumental learning involves the dynamic interaction of a computationally rigid, low-level striatal and a more flexible, high-level prefrontal component.
Methods: To evaluate the role of the prefrontal cortex in instrumental learning, we applied anodal transcranial direct current stimulation (tDCS) optimized for the left dorsolateral prefrontal cortex, by using realistic MR-derived finite element model-based electric field simulations. In a study with a double-blind, sham-controlled, repeated-measures design, sixteen male participants performed a probabilistic learning task while receiving anodal and sham tDCS in a counterbalanced order.
Results: Compared to sham tDCS, anodal tDCS significantly increased the amount of maladaptive shifting behavior after optimal outcomes during learning when reward probabilities were highly dissociable. Derived parameters of the Q-learning computational model further revealed a significantly increased model parameter that was sensitive to random action selection in the anodal compared to the sham tDCS session, whereas the learning rate parameter was not influenced significantly by tDCS.
Conclusion: These results congruently indicate that prefrontal tDCS during instrumental learning increased randomness of choice, possibly reflecting the influence of the cognitive prefrontal component.

1. Introduction

In most everyday situations, we constantly have to adapt and optimize our behavior to cope with various, often conflicting, demands and constraints posed by each specific environment. An important aspect of adaptive behavior is the capability of choosing those actions that lead to a high amount of cumulative reward. One way to achieve this goal is by successively

* *Corresponding author*. Department Clinical Neurophysiology, University Medical Center, Georg-August University, Robert-Koch-Str. 40, Göttingen D-37075, Germany.
E-mail address: zsoltturi@gmail.com (Z. Turi).
[1] These authors contributed equally.
http://dx.doi.org/10.1016/j.cortex.2014.08.026

generating predictions about the consequences of each action. Generating and using these predictions to guide behavior is known as instrumental learning (Dayan & Balleine, 2002).

Instrumental learning in humans recruits multiple, functionally interacting and parallel brain systems (for reviews see Dolan & Dayan, 2013; Samson, Frank, & Fellous, 2010); these involve a striatal reinforcement learning (RL) component and a cognitive, prefrontal control component (Collins & Frank, 2012; Daw, Niv, & Dayan, 2005; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006), also known respectively as the model-free and model-based controls of instrumental learning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Daw et al., 2005; Wunderlich, Smittenaar, & Dolan, 2012). The low-level RL (or model-free) component is characterized by computational rigidity and it requires a large number of learning trials to gradually integrate the long-term probability of reinforcement values in response to probabilistic reward associations (Frank, Moustafa, Haughey, Curran, & Hutchison, 2007).

The high-level cognitive (or model-based) component, driven by the prefrontal system, has greater computational flexibility as it dynamically computes the policy to optimize behavior by evaluating the instrumental requirements of the decision situation (Daw et al., 2006). On the one hand, this is achieved by actively maintaining the reinforcement history in working memory (WM) which permits fast goal-directed decisions, albeit with the restriction of a limited capacity (Collins & Frank, 2012; Frank, et al., 2007). On the other hand, functional neuroimaging evidence also suggests that the prefrontal system controls adaptive exploration (Daw et al., 2006). Further evidence also indicates the role of prefrontal involvement specifically, as individual genetic differences in regulating prefrontal dopamine (DA) Catechol-O-methyltransferase (COMT) rs4680 single nucleotide polymorphism has an impact on exploratory behavior but not on the level of striatal DA (Frank, Doll, Oas-Terpstra, & Moreno, 2009).

Nevertheless, genetic studies are correlational in nature and a more direct demonstration of the involvement of the prefrontal component in cognitive control in instrumental learning requires a focal interference with prefrontal regions. Transcranial direct current stimulation (tDCS) has the potential to temporarily shift neuronal membrane potentials of a given neuronal population by passing a low-intensity electrical current through the brain (Nitsche & Paulus, 2000). These physiological effects have been linked to changes in a wide range of cognitive functions, including those that are related to the prefrontal cortex, such as WM (e.g., Zaehle, Sandmann, Thorne, Jäncke, & Herrmann, 2011) or prototype learning (Ambrus et al., 2011).

Modeling studies investigating the tDCS-induced current profile characteristics indicate that the effect of tDCS, at least from electrodes in close spatial proximity, is primarily limited to the neocortex (Datta, Elwassif, Battaglia, & Bikson, 2008; Faria, Hallett, & Miranda, 2011), although tDCS may have the ability to remotely activate deeper brain structures, such as the striatal system (Chib, Yun, Takahashi, & Shimojo, 2013). The common notion that anodal tDCS leads to an increase and cathodal tDCS leads to a decrease in neuronal excitability in the brain area underneath the electrode have been challenged by recent evidence (Reato et al., 2013). First, the electric field induced by tDCS can both de- and hyperpolarize within the same gyrus (Reato et al., 2013) and second, different types of neurons are differentially modulated depending on their morphology and axonal orientation (Radman, Ramos, Brumberg, & Bikson, 2009). Hence, a simple mechanistic relation between polarity and expected behavioral changes may be difficult to establish. Indeed, recent evidence suggests that tDCS has less consistency in polarity effects in cognitive tasks compared to basic motor functions (Jacobson, Koslowsky, & Lavidor, 2012).

The aim of the present experimental work has been to study, which component of instrumental learning was influenced by prefrontal tDCS by evaluating the effect of anodal tDCS on behavior as measured by accuracy and computational model parameters. Advances in computational modeling of RL using Q-learning algorithms allow distinct processes to be modeled in detail. This entails the ability to derive information about how performance is affected by specific behavioral influences or strategies by fitting the RL model to behavioral data (e.g., Frank et al., 2009).

In the classical model we employed in this study (Jocham, Klein, & Ullsperger, 2011), the learning rate parameter α reflects the impact of the prediction error (i.e., the difference between the previous outcome estimate and the actual estimate after a certain action). Larger α values reflect trial-to-trial fluctuations (a recency effect), whereas lower values indicate a gradual value integration and more stable value estimation (Frank et al., 2007). If prefrontal anodal tDCS biases participants to rely more on the WM component, we expected to observe a trial-to-trial behavioral adjustment (i.e., change of decision after negative response) during learning and an increased α value. In contrast, if anodal tDCS compels participants to rely less on the WM component, then a lower α value and less trial-to-trial behavioral adjustment will be observed – which would increase outcome-dependent exploitation of the better symbol. In addition, the β parameter, also known as the "temperature" or "noise" parameter, reflects the learners' bias towards either exploitation (i.e., choosing the better option in case of lower β values) or exploration (i.e., choosing the items more randomly; higher β values) (Frank et al., 2007; Jocham, et al., 2011). This model is designed to capture behavior in a probabilistic environment where not only the expected value (determined by integrating past outcomes with learning rate α) determines the decision, but choices are also characterized by intrinsic randomness, reflected in the noise parameter β (Beeler, Daw, Frazier, & Zhuang, 2010). If anodal tDCS affects exploration and induces randomness in choices, participants will demonstrate increased shifting behavior (i.e., a tendency to change, rather than repeat a response to the same stimulus) and a decreased preference for symbols that are associated with the higher reward probability, reflected by higher β values.

2. Material and methods

2.1. Participants

Sixteen right-handed, healthy, native German-speaking participants took part in the study (mean age of 22.9 ± 2.2 years). In order to avoid menstrual cycle-dependent level changes of