Cortical hemodynamic mechanisms of reversal learning using high-resolution functional near-infrared spectroscopy: A pilot study

Charlotte Piau, Mahdi Mahmoudzadeh, Astrid Kibleur, Mircea Polosan, Olivier David, Fabrice Wallois

Version of Record: https://www.sciencedirect.com/science/article/pii/S0987705321000861
Manuscript_c61bc3bae9fed7e926303b99860a1f6e

Original article

Cortical hemodynamic mechanisms of reversal learning using high-resolution functional near-infrared spectroscopy: A pilot study

Authors:

Charlotte Piau[1], Mahdi Mahmoudzadeh[1-2], Astrid Kibleur[3], Mircea Polosan[3], Olivier David[3-4], Fabrice Wallois[1-2]

Affiliation:

[1]INSERM U1105, Université de Picardie, GRAMFC, CURS, CHU Sud, rue René Laennec, 80036 Amiens Cedex 1, France

[2] INSERM U1105, EFSN Pédiatrique, CHU Amiens sud, Avenue Laennec, 80054 Amiens cedex

[3] Univ Grenoble Alpes, INSERM U1216, GIN, Grenoble Institut des Neurosciences, 38000 Grenoble, France

[4] Aix Marseille Univ, INSERM U1106, INS, Institut de Neurosciences des Systèmes, 13005 Marseille, France

Running title:

fNIRS imaging in reversal-learning tasks

Abstract

Objectives: Reversal learning is widely used to analyze cognitive flexibility and characterize behavioral abnormalities associated with impulsivity and disinhibition. Recent studies using fMRI have focused on regions involved in reversal learning with negative and positive reinforcers. Although the frontal cortex has been consistently implicated in reversal learning, few studies have focused on whether reward and punishment may have different effects on lateral frontal structures in these tasks.

Methods: During this pilot study on eight healthy subjects, we used functional near infra-red spectroscopy (fNIRS) to characterize brain activity dynamics and differentiate the involvement of frontal structures in learning driven by reward and punishment.

Results: We observed functional hemispheric asymmetries between punishment and reward processing by fNIRS following reversal of a learned rule. Moreover, the left dorsolateral prefrontal cortex (l-DLPFC) and inferior frontal gyrus (IFG) were activated under the reward condition only, whereas the orbito-frontal cortex (OFC) was significantly activated under the punishment condition, with a tendency towards activation for the right cortical hemisphere (r-DLPFC and r-IFG). Our results are compatible with the suggestion that the DLPFC is involved in the detection of contingency change. We propose a new representation for reward and punishment, with left lateralization for the reward process.

Conclusions: The results of this pilot study provide insights into the indirect neural mechanisms of reversal learning and behavioral flexibility and confirm the use of fNIRS imaging in reversal-learning tasks as a translational strategy, particularly in subjects who cannot undergo fMRI recordings.

Keywords: functional neuroimaging, fNIRS, reversal learning, neurovascular coupling, cognitive flexibility, reward, punishment

Introduction

Humans must be able to adapt to changes in their environment. This requires quickly adjusted responses to voluntarily inhibit or alter established behavior (prepotent response) [45]. Paradigms such as reversal-learning tasks (RLTs) can be used to measure behavioral flexibility [20*;* 45*;* 60*;* 121]. The RLT paradigm provides an approach to measure a participant's capacity to select an appropriate behavior (i.e., response) when the rules of the environment are modified. For example, participants must first choose one stimulus (e.g., picture or action) associated with the desired outcome (to win money, for example, as positive feedback). Then, there is an alteration of the rule, and the stimulus associated with the positive feedback changes, a reversal occurs, and participants must select the new correct stimulus related to the desired outcome to appropriately update the response [45]. Impairments in reversal-learning processes are associated with a wide range of abnormal behavioral, neurological and psychiatric conditions characterized by impulsiveness and disinhibition, such as obsessive compulsive disorder (OCD) [20], conduct disorder [14], bipolar disorder [38], Parkinson disease (PD) [24], and major depression [93]. More specifically, impairments related to poor reversal learning are associated with dysregulation and poor dopamine (D2 receptor) modulation of the frontocorticostriatal circuitry [59]. For example, certain unmedicated PD patients have shown altered learning relative to medicated PD patients [26; 43]. Such a lack of cognitive flexibility was shown to be associated with maladaptive patterns of repetitive, inflexible cognition and behavior in OCD patients [46].

Prefrontal lesional studies on monkeys [15; 37; 58] and humans [41; 55] have been performed to characterize the underlying neural mechanisms that are solicited in the RLT paradigm) and have consistently reported involvement of the ventrolateral prefrontal cortex (VLPFC) and lateral orbitofrontal cortex (OFC). Thus, Fellows et al. [41] studied the ventromedial prefrontal cortex vs DLPFC of subjects in three subgroups (8 to 12 people, 56-62 years old, 50-60% women) free of neurological or psychiatric disease and not taking psychoactive medications and excluded subjects with past or intercurrent medical or neurological disease likely to impair cognition. Moreover, Hornak et al. [55] studied 31 patients divided into five sub-groups (3 to 6 people, 19-72 years old, 20-100% male). They excluded subjects with damage outside of the prefrontal cortex, those with an alcohol- or drug-dependence, and those with a full-scale IQ below a cut-off of 80. In addition, functional magnetic resonance imaging (fMRI) studies have highlighted the involvement of other brain regions, including the lateral OFC [20; 48; 83], inferior frontal gyrus (IFG) [25;80], dorsomedial prefrontal cortex

(DMPFC) [14], dorsolateral prefrontal cortex (DLPFC) [92; 45], posterior parietal cortex [47], anterior cingulate cortex (ACC) [21; 60; 64], and striatum [25; 49].

A number of pharmacological and functional imaging studies have reported distinct mechanisms underlying learning from positive and negative feedback [26; 27; 43]. Reward and punishment represent two major motivations in learning in a changing environment and the adaptation of behavior. However, how reward and punishment may modulate reversal learning has been less frequently examined [121]. Concerning the decision-making process, Reber and Temel's reviews [91] reported social, emotional, and decision-making impairments not only associated with unilateral wmPFC damage [4; 98] but also with neurological lesions in other structures [22].

In terms of punishment learning, Wheeler & Fellows [118] found that the vmPFC (that includes the mesial part of OFC) is specifically involved in learning by negative feedback during a reversal learning task. In 2012, Palminteri et al. [87] suggested the existence of a distinct punishment-based avoidance learning system based on the results of a Go/No Go task associated with damage to the insula or dorsal striatum. The use of fMRI during nonclassical standard reversal paradigms [94], in which the received outcome does not depend on the subject's choice, has shown that the posterior dorsal striatum responds only to unexpected rewards, whereas the anterior ventral striatum responds to both unexpected rewards and unexpected punishment. However, these lesion studies [118;41] did not directly compare reversal learning by reward and punishment. Thus, it is unclear whether similar dissociations between reward and punishment exist for reversal learning.

Several RLT studies have reported asymmetry in activated regions, with involvement of the left orbitofrontal hemisphere in reward feedback [83;84; 92], whereas the right frontal hemisphere regions [right OFC: 83; right DLPFC: 121] showed greater sensitivity to punishment. Studies on tasks other than RLT have reported frontal lateralization depending on the valence of the feedback, in which positive outcomes were lateralized to the left frontal regions [12; 105;120] and unpleasant feedback to the right frontal hemisphere.

Thus, it is important to understand how and where reward and punishment modulate reversal learning and how these learning processes have been affected in conditions associated with a lack of flexibility. Recently, OCD patients have been shown to benefit from deep brain stimulation, similarly to PD patients [19]. However, the spatial extension of these effects cannot be evaluated, as these patients with deep brain electrodes usually cannot be recorded by fMRI. We propose to use an alternative tool, functional near infrared spectroscopy

(fNIRS), which shows good spatial resolution when a sufficient number of optodes are used [119].

To date, the RLT paradigm has not previously been studied using this approach. Replicating the fMRI findings using fNIRS is of particular translational value for future studies on subjects who cannot be recorded by fMRI, (e.g., subjects with brain electrodes, pacemakers, claustrophobia, and other contraindications).

fNIRS is used non-invasively in a natural environment for human infants and adults to analyze cortical activation [71; 74; 119], as well as in pathological situations [95]. Brain activation is associated with neurovascular coupling that induces hemodynamic changes, modifying the optical properties of brain tissue, which can be assessed by fNIRS [62; 71]. In this pilot study, we used the more powerful temporal resolution of fNIRS [30], relative to fMRI, to investigate the cortical hemodynamic response to neuronal activation by analyzing the changes in oxygenated and deoxygenated hemoglobin concentrations (HbO and HbR) induced by RLT. We expected to learn about the brain regions involved in rightward activation, at least in the DLPFC, under the punishment condition and leftward activation asymmetry under the reward condition during RLT. Moreover, we aimed to investigate the involvement of the DLPFC, IFG, and OFC in RLT, as observed by fMRI [25; 45; 60; 80; 92; 121]. Thus, we aimed to (1) provide an estimate of the involvement of the DLPFC and IFG during the reversal process and, (2) differentiate the temporal and spatial characteristics of the responses under reward and punishment learning conditions using an innovative noninvasive portable solution based on high density fNIRS due to its good spatial resolution.

Methods

Participants

Eight healthy right-handed French native-speaking participants (6 women and 2 men) aged between 22 and 56 years were enrolled in this pilot study and were not paid. All subjects had normal or corrected to normal vision, followed or had followed a graduate school degree, and had no history of neurological disease. They were asked to sit comfortably and limit their head movements as much as possible while performing the experiment. The experiment was conducted in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the protocol of the study was approved by the local ethics committee and the Comité de la Protection des Personnes (NO II N°2013-A01297-38).

Informed written consent was obtained from each subject before the experiment. Subjects were informed that the task was based on the hypothetical gain/loss of money.

The deterministic reversal learning task

During this pilot study, participants sat in a chair 30 cm from a 23-inch computer monitor in a dark room. They also had a tablet with a keyboard. The task was composed of 400 stimuli (including 80 trials for the training blocks) and lasted approximately 25 mins. It was divided into six blocks (2 training blocks followed by 4 task blocks). The two training blocks consisted of one positive and one negative training block with 40 positive and 40 negative stimuli, respectively. Each of the two training blocks had a duration of 180 s. After the two training blocks, the task was continued with four "process" blocks (*Figure 1A*). Each process block had a duration of 360 s and was composed of 80 randomized trials (40 positive and 40 negative). The organization of the task is shown in *Figure 1*. Every trial began with presentation of a pair of animal pictures (e.g., the stimuli) from the IAPS [69], of neutral valence and intensity, presented on a grey background for a maximum of 3000 ms. Each presentation was followed by a short feedback followed by a fixation-cross interstimulus lasting for 1000 to 1500 ms (randomly determined). The positions of the pictures of the animals were randomized (one on top, the other at the bottom of the screen). During the 3000 ms, participants had to decide behind which animal money was hidden. Once the response was made by the participant, feedback was delivered (*Figure 1C*). There were two types of trials: *positive feedback pairs* and *negative feedback pairs*. For the positive feedback pairs, one picture was associated with symbolic monetary gain feedback (+100$) and the other with neutral feedback (0$). For the negative feedback pairs, one picture was associated with symbolic monetary loss feedback (-100$) and the other with neutral feedback (0$). The order of presentation of the positive and negative trials was randomized. Reversal occurred after four to six correct answers for both the positive and negative feedback trials. In other words, after a certain period of time, the picture associated with a reward in the positive pair was associated with neutral feedback and the other picture from the positive pair was associated with the reward. Similarly, in the negative pair, the picture associated with the punishment was associated with the neutral feedback and the other picture from the negative pair was then associated with the punishment. Thus, the rules established during the acquisition phase (A) (phase of 4 to 6 trials with the same rule, during which the participants learn the stimuli/feedback associations) are reversed during the reversal phase (R). After 4 to 6 correct responses (CR) for the same pair during the reversal phase, a new reversal occurred. The number of correct consecutive answers (4 to 6) necessary for the switch was randomized. The

picture of the animal was changed between odd and even blocks to avoid learning the rule of the association between the animals and the feedback. The task was programmed using E-Prime software and synchronized with the fNIRS system using an external trigger that was sent by E-prime to the fNIRS system to identify when each acquisition and reversal stimulus was presented to the subject.

[Insert Figure 1 about here]

Behavioral responses

Behavioral responses were labelled according to Reminsje et al. [92] (*Figure 1B*). During the acquisition phase, the correct answers (ACR) and erroneous answers (AE) were differentiated. During the reversal phase, the correct answers (RCR) and reversal errors (RE) were also identified following each switch of the rules. Among all reversal errors (RE) following the switch, the last error just prior to learning the new rule (*i.e.,* the first CR) was called the '*final reversal error (FRE)'*. After the first CR, a false response, the '*no link to the switch (RENS)'* was considered as the participant having learned the new rule.

Data acquisition/analysis

fNIRS Recording

fNIRS signals were recorded with a portable continuous wave MEDELOPT® system (Seenel Imaging TM) using 32-detectors and 16-emitters, resulting in 512 potential channels (*Figure 2E*). Optodes covered the prefrontal, frontal, and parietal lobes, with an emitter-detector distance of 2.5 cm (*Figure 2A*). The sampling rate was 2 Hz (500 ms each sample) and data were recorded at two wavelengths (660 and 850 nm). The position of the headgear was checked before and after the experiment; photos were taken to review placement and optode positions were digitalized for each subject using a 3D digitizer (NDI Medical Polaris Vega TM). No subject was excluded for an incorrect fNIRS sensor location. Optode positions (sources and detectors) were defined according to the EEG 10-10 system coordinates to standardize the headgear position among the participants. The lower edge of electrode positions was located over the frontal area, with detector 2 (D2) centered above the highest point of the eyebrow (Fp2) (*Figure 2B*). The headgear covered the temporal area, with detector 27 (D27) above the C line (C1).

The sensitivity of various configurations was assessed using the AtlasViewer toolbox of Matlab to evaluate the best configuration of source and detector combinations for scanning the frontal and temporal lobes [2] (*Figure 2A and 2B*). The fNIRS sensitivity map (*Figure*

2C) was modelled using AtlasViewer and freely available Monte-Carlo photon transport software (tMCimg routine, number of simulated photons = 10^6). The optimal sensitivity configuration of the fNIRS source and detector positions is presented in *Figure 2A*.

[Insert Figure 2 about here]

Data analysis

The response of the participants associated with detection of the unexpected outcome for each trial following a switch of the parameters was assessed by the a*ccuracy* (the percent of correct answers) and *reaction time* (the time, in seconds, used by the participant to choose the animals). Each presentation following a switch was ordered for each trial: the first presentation following a new rule was designated 1, the second presentation 2, etc., until the new switch of rules. The average accuracy and reaction time were then computed for each numbered presentation. The various parameters were compared between the first and following presentations for each condition to analyze the effect of repetition. Difficulties in updating the strategy were indicated by successive errors (*perseverative reversal errors)* and occurred when the former discrimination strategy, which was now obsolete, was still used.

Several studies have evaluated the effect of reward and punishment feedback on reversal learning by focusing on contrasting reversal learning with the acquisition phase [14; 45; 92; 121]. Hemodynamic analysis was performed according to Remijnse [92]. The reversal effect (*i.e.*, unexpected outcome) was then characterized by contrasting the responses recorded during reversal learning and those during the acquisition phase: (FRE+RE) – (AE+RENS) (*Figure 1*). Such a contrast was performed here for the reward and punishment conditions. Contrasts were performed at the single-subject level to account for individual characteristics.

Homer2 Matlab toolbox was used to analyze the fNIRS signal [57]. A band-pass filter between 0.03 and 0.1 Hz was applied to eliminate physiological noise (very low-frequency oscillations, respiration, and heartbeat). To determine periods of motion artifacts in the fNIRS datasets, we also used a motion artifact identification algorithm (i.e. hmrMotionArtifact.m) integrated into the HOMER2 NIRS processing package (http://www.homer-fnirs.org). This algorithm provides reliable identification of motion artifacts based on changes in signal amplitude and/or standard deviation and is similar to the approach described by Scholkmann et al. [100] and Cooper et al. [28]. Moreover, to maximize the accuracy of motion identification, we visually inspected the results and manually rejected any motion artifacts. This algorithm is not channel-specific, i.e., signal changes that are determined to be artifacts

in one channel are marked as motion in all channels; this relies on the reasonable assumption that motion artifacts affect multiple channels. A window from -5 to 30 s around the onset of the stimulation (t = 0 s, when a pair of animal pictures was presented) was used to analyze the hemodynamic response to RLT [16]. The fixation-cross interstimulus interval throughout the entire task served as the implicit baseline to allow estimation of the hemodynamic response function by simple averaging [29].

The weighted arithmetic average cerebral hemodynamic response was then computed for each type of stimulation for all combined blocks and each subject. The data were z-score normalized for each subject to harmonize the data according to the individual characteristics. Then, a baseline correction [-5, 0] s was finally applied to the normalized data.

Regions of interest (ROI)

We were able to associate each channel with a cortical ROI using the AtlasViewer toolbox [2] based on the digitalized coordinates recorded with a 3D digitizer (Polaris Vega). Based on the MNI coordinates of each optode, our setup covers seven ROI: the right prefrontal cortex superior (r-PFC), right inferior frontal gyrus (r-IFG), right dorsolateral prefrontal cortex (r-DLPFC), left dorsolateral prefrontal cortex (l-DLPFC), left prefrontal cortex superior (l-PFC), left inferior frontal cortex (l-IFG), and orbitofrontal cortex (OFC) (*Figure 2D*). Channels with an inter-optode distance < 5 cm were then selected [71]. Several studies previously tested various inter-optode distances to optimize the sensitivity of the signal due to the dispersion properties of the photons. They observed that the relative contribution of extracranial tissue decreased as the inter-optode distance increased. Thus, we focused on an inter-optode distance < 5 cm [104] and kept 128 relevant channels (OFC: 8 channels, r-DLPFC: 28 channels, l-DLPFC: 20 channels, r-IFG: 22 channels, l-IFG: 20 channels, r-PFC: 8 channels, and l-PFC: 21 channels).

The cerebral hemodynamic response per ROI was computed for each subject by averaging the responses of channels covering the same ROI. Finally, we computed the grand average per ROI by averaging the cerebral hemodynamic response of all subjects for the same ROI. The shape and peaks of the curves were analyzed to compare the cerebral hemodynamic responses between ROIs under each condition. Three other parameters were also analyzed for various sub-periods from 0 to 30 s (every 5 and 10 s): the average, slope, and area under the curve (AUC).

Thus, the *'average'* of the Hb values (HbO, HbR, or HbT) was computed for each timepoint of each subperiod. The dynamics of the responses were characterized by calculating the *'slope'* for each considered subperiod between t0 and the first maximum absolute amplitude

(i.e. either peaks or valleys). We also examined the slope coefficient, which is indicative of the magnitude (and the direction) of the oxygenation responses over the stimulation period. Thus, a higher (and positive) slope value for HbO is associated with greater and faster cortical activation [76]. The cerebral hemodynamic response over each region was characterized by computing the power of the activation, that is the cumulative sum of each point of the average (AUC) from the beginning to the end of the considered subperiod. Finally, the temporal evolution of the activation was visualized by projecting the cerebral hemodynamic response on the cortex for the selected channels using the AtlasViewer Matlab toolbox for HbO and HbR for the various subperiods (i.e. [-5 0], [0 5], …, [25 30] s) [2].

Statistical Analysis

We investigated the differences in the dynamics of the response between the two conditions across various sub-periods (i.e. [-5 0], [0 5], …, [25 30] s) to assess the changes in concentration. The normality of the data was verified using the Shapiro Wilk normality test, adapted to the small number of participants, allowing us to use the student t-test. Two-way variance analysis (ANOVA) was used to investigate the hemodynamic response and the effect of the Condition*ROI on modifications for the response of HbO, HbT, and HbR across the various subperiods. In addition, Fisher post-hoc multiple comparison analyses were performed to identify which ROI and which condition had a significative effect on Hb. Moreover, we investigated the significant hemodynamic changes, using t-tests, between each 500 ms acquisition point after stimulation ([0, 30] s) for each ROI relative to the average hemodynamic concentration during the baseline ([-5, 0] s) for the same ROI. Then, we tested for significant differences for the three parameters (average, slope, and AUC) between the reward and punishment conditions across the various subperiods using student t-test statistics. For pilot studies, Lee et al. recommend extending the confidence interval, for example to 85% or 75% [70]. We report both the significant results ($p < 0.05$) and the tendencies ($p < 0.09$) to focus on the impact of reversal processing.

Results

Behavioral analysis

The behavioral data for each repetition (1-8) of pairs on the screen during the acquisition and reversal phase for both the reward and punishment conditions are presented in *Figure 3*. We compared the accuracy and reaction time parameters. During the acquisition phase (*Figure 3A – left panel*), between the first and following presentations (2-8), the accuracy increased non-

significantly (*Figure 3A - left panel*) for the reward (accuracy of presentation 1 vs 2-8: +17%, t_{value} = -1.22, p = 0.26) and punishment (accuracy of presentation 1 vs 2-8: +29%, t_{value} = -2.37, p = 0.05) conditions. We observed no significant differences for the averaged accuracy between the punishment and reward conditions (t_{value} = -0.82, p = 0.4).

The reaction time (RT) (*Figure 3B – left panel*) decreased significantly between the first and second presentation under both the punishment (RT of presentation 1 vs 2: -0.4 s, (t_{value} = 2.98, p = 0.02) and reward (RT o presentation 1 vs 2: -0.5 s, t_{value} = 4.18, p = 0.01) conditions. We observed no significant differences between the two conditions (t_{value} = -0.56, p = 0.6) for the averaged RT.

During the reversal phase (*Figure 3A – right panel*), between the first and following presentations (2-8), the accuracy (*Figure 3A – right panel*) increased to 69% for the punishment (presentation 1 vs 2-8, t_{value} = -9.28, p = 0.001) and 57% for the reward conditions (presentation 1 vs 2-8 t_{value} = -5.46, p = 0.001), suggesting that the participants did not predict the reversal. More precisely, after the first presentation, the subjects were able to immediately reconfigure the stimulus-reward association upon the second presentation (accuracy of presentation 1 vs 2: +46% for the reward [t_{value} = -4.63, p = 0.002] and 66% for the punishment [t_{value} = -8.22, p = 0.001] conditions, respectively). The differences for the averaged accuracy between the reward and punishment conditions ($t_{valu\ e}$= -7.27, p = 0.01) were significant for all presentations. For the RT (*Figure 3B – right panel)*, we observed a non-significant increase between presentation 1 and presentation 2 for the punishment (-0,06 s, t_{value} =- 1.33, p = 0.23) and reward (-0.08 s, $t_{valu\ e}$= -2.05, p = 0.08) conditions. Comparison of the averaged reaction time for all presentations showed the difference between the two conditions to be significant (t_{value} = -11.07, p = 0.001).

Finally, the increase in the number of perseverative errors (*Figure 3C*) between the reward and punishment conditions was not significant (t_{value} = -2.06, p = 0.07), suggesting that there may be strong cognitive costs when expressing the new associations under the reward condition. Based on the significantly lower accuracy, as well as the tendency towards an increase in the number of perseverative errors, these behavioral results suggest that the reversal may more negatively affected the reward condition than the punishment condition.

[Insert Figure 3 about here]

Cortical hemodynamic response

We analyzed the detection of the unexpected outcome by computing the hemodynamic event contrast as described by Remijnse, [92]: (FRE+RE)-(AE+RENS). Cortical activation from fNIRS signals for an expected outcome is characterized by a substantial increase in HbO, with a lower delayed decrease in HbR [42; 89].

Brain regions differentially involved in reversal under punishment and reward conditions

We determined which ROIs were activated under the two conditions by investigating significant activation relative to baseline ([-5, 0] s). We observed significant activation for the left hemisphere for both conditions (*Figure 4A*). There was no significant hemodynamic response in the right hemisphere. Under the reward condition, l-DLPFC activation consisted of a hemodynamic response characterized by a significant increase in HbO (p = 0.008) on $t_{04.5\text{-}16.5}$, with a latency to the peak of approximately 7 s (t_{value} = 4.41, p = 0.003, 1.35 AU) and a slope for the coefficient of c = 0.13. The change in HbR was inverted, smaller, and nonsignificant. The l-IFG showed a significant HbO response (p = 0.03) on $t_{02\text{-}07}$ and $t_{11.5\text{-}13.5}$, with a similar pattern and latency to the peak (approximately 3 s, 1.3 AU, c = 0.21, t_{value} = 2.85, p = 0.025). There was a second peak at approximately 12 s (0.8 AU, t_{value} = 2.69, p = 0.031). Under the punishment condition, the hemodynamic response in the right hemisphere was not significant but we observed a tendency of a hemodynamic activation pattern. Thus, the hemodynamic responses of the r-DLPFC r-IFG consisted of an increase in HbO (r-DLPFC: 1.43 AU, c = 0.34, with t = [4.5, 6] s, t_{value} = 2.15 p = 0.07; r-IFG: 1.09AU, c = 0.20, with t = [2.5, 4] s, t_{value} =2.14, p = 0.07), with a latency to the peak of approximately 5 s and a decrease in HbR. The r-IFG showed a second peak at approximately 13.5 s (1.42 AU, with t = [12, 15.5] s, t_{value} = 2.2, p = 0.07). Finally, the OFC showed a significative increase in HbO, with a latency to the peak of approximately 13 s (1.45 AU, c = 0.11, t_{value} = 3.47, p = 0.01).

[Insert Figure 4 about here]

Comparison of the AUC for the 5-s subperiods to that of the baseline showed involvement of the l-DLPFC (t_{value} = 4.54, p = 0.002 $t_{05\text{-}10}$ and t_{value} = 5.38, p = 0.001 $t_{10\text{-}15}$ and l-IFG (t_{value} =2.76, p = 0.03 $t_{00\text{-}05}$; t_{value} = 2.68, p = 0.03 $t_{05\text{-}10;}$ and t_{value} = 2.70, p = 0.03 $t_{10\text{-}15}$) under the reward condition. Under the punishment condition, we also observed the involvement of the OFC (t_{value} = 2.64, p = 0.03 $t_{00\text{-}05}$; t_{value} = 5.69, p = 0.0007 $t_{05\text{-}10;}$ and t_{value} = 4.11, p = 0.005 $t_{10\text{-}15}$) and a tendency towards involvement of the r-DLPFC (t_{value} = 2.06, p =0.08 for $t_{03\text{-}08}$) and r-IFG (t_{value} = 2.10, p = 0.074 for $t_{00\text{-}05}$ and t_{value} = 2.07, p = 0.077 on $t_{10\text{-}15}$).

We then evaluated these differences in activation between the two conditions (reversal vs punishment) for the same ROI, considering the three parameters (slope, AUC, average) across the subperiods.

Statistical analysis (*Figure 4B*) of the changes in HbO in the l-DLPFC showed a significant difference between conditions for the AUC ($t_{value} = 3.51$, $p = 0.01$) and average ($t_{value} = 3.26$, $p = 0.013$) for the period $t_{10\text{-}15}$. There were also significant differences for HbT (AUC, $t_{value} = 2.84$, $p = 0.025$; average, $t_{value} = 2.49$, $p = 0.04$). We observed no significant differences in any parameters for HbR. For the r-DLPFC, we observed significant differences between conditions for HbO (slope, $t_{value} = -2.79$, $p = 0.02$ on $t_{00\text{-}15}$) (*Figure 4C*) and HbT (slope, $t_{value} = -2.92$, $p = 0.03$ on $t_{00\text{-}15}$), but none for HbR or any other parameters.

The r-IFG showed significant differences between conditions for HbR for $t_{00\text{-}10}$ (slope: $t_{value} = 2.98$, $p = 0.020$; AUC: $t_{value} = 2.05$, $p = 0.080$; average: $t_{value} = 1.94$, $p = 0.09$), but none for HbO or HbT (*Figure 4F)*. The l-IFG showed a tendency for HbO for $t_{00\text{-}05}$ (average: $t_{value} = 2.22$, $p = 0.06$; AUC: $t_{value} = 2.39$, $p = 0.05$) (*Figure 4E*). We also observed significant differences for the average and AUC for various subperiods ($t_{05\text{-}15}$. $t_{05\text{-}10}$, $t_{10\text{-}20}$, $t_{10\text{-}15}$, and $t_{00\text{-}20}$) for HbO and HbT in the OFC ($t_{value} = -2.78$, $p = 0.04$) (*Figure 4D*).

Brain regions differentially affected by reversal by the effect of ROI × condition

We evaluated the effect of the condition and ROI on the Hb value by performing ANOVA on *Condition (Positive, Negative)* × *ROIs (OFC, R-PFC, R-DLPC, R-IFG, L-PFC, L-DPLFC, L-IFG)* for the three previously identified significant subperiods (*Table S1*). We observed a tendency towards a significant effect of the *ROI* × *condition* on HbO (Average $t_{05\text{-}15}$ $F_{(6,98)} = 1.93$, $p = 0.074$; AUC $t_{05\text{-}15}$ $F_{(6,98)} = 2.00$, $p = 0.072$) for $t_{05\text{-}15}$. We also observed a tendency towards a significant effect of *ROI* × *condition* (*Table S1.A*) for the HbR (slope $t_{00\text{-}15}$ $F_{(1,98)} = 1.92$, $p = 0.085$) and a significant effect of the *condition* for HbO (Average $t_{10\text{-}20}$ $F_{(1,98)} = 5.26$, $p = 0.024$; AUC $t_{10\text{-}20}$ $F_{(1,98)} = 4.53$, $p = 0.035$) and HbT (AUC $t_{10\text{-}20}$ $F_{(1,98)} = 5.88$, $p = 0.017$; Average $t_{10\text{-}20}$ $F_{(1,98)} = 6.25$, $p = 0.014$; Slope $t_{00\text{-}15}$ $F_{(1,98)} = 3.23$, $p = 0.076$). An effect of the *ROI* (*Table S1.B*) was observed for the HbO slope parameters (slope $t_{00\text{-}15}$ $F_{(6,98)} = 19.32$, $p = 0.001$).

We then performed a Fisher post-hoc analysis to define the ROIs and conditions that showed a significant effect. The analysis was first performed for the $t_{05\text{-}15}$ period and if no significant results were observed, it was performed for the $t_{05\text{-}10}$ and $t_{10\text{-}15}$ periods. For clarity, we present the results in three cross-tables, one for each parameter (slope, AUC, and average) (*Figure 5A*). A cross-table summarizes the pairwise comparison of results from a multiple combination test. Thus, we determined whether there was a difference between the column

ROI and the row ROI. Thus, two ROIs which were significantly different are symbolized by one or several asterisks (*), the number of asterisks depending on the *p*-value (*: < 0.05, **: < 0.01, ***: < 0.001, and *t* for tendency < 0.09). The ROI were organized by condition (reward and punishment) and then hemisphere (right, left, and both for the OFC). For example, the HbO matrix (*Figure 5A*) for the average parameter showed a significant difference between the right and left DLPFC under the reward condition and for three ROIs under the punishment condition: OFC and a tendency towards a significant difference for the r-DLPFC and r-IFG. These observations confirm the significant difference between the right and left DLPFC under reward conditions (average, $p = 0.04$; AUC, $p = 0.05$) and the OFC between the two conditions (AUC, $p = 0.02$; average, $p = 0.04$). There was also a significant difference for the slope parameters between the two conditions for the r-DLPFC and r-IFG (r-DLPFC : $p = 0.03$; r-IFG : $p=0.05$).

The statistical *post-hoc* analysis is added to the HbO hemodynamic mapping in *figure 5B*. The post-hoc statistical analysis for the average parameters confirmed the first observations: DLPFC activation was significantly stronger ($p = 0.04$) on the left than right hemisphere at t_{00-10} under the reward condition. Comparison of the two conditions showed a tendency for the r-DLPFC to be less activated under the punishment condition than the reward condition at t_{00-10} ($p = 0.065$). This last result was also observed for the r-DLPFC under the punishment condition compared to the r-IFG under the reward condition ($p = 0.078$).

The fNIRS data showed significant results for the reward condition for the left hemisphere and involvement of the OFC and a tendency towards involvement of the right hemisphere under the punishment condition (*Figure 5C*). The involvement of the l-IFG and l-DLPFC under the reward condition and the OFC under the punishment condition showed the usual hemodynamic pattern, characterized by an increase in [HbO] and a decrease in [HbR]. Over the right hemisphere, only a tendency ($p =0.07$) towards a typical hemodynamic response was observed over the r-DLPFC and r-IFG. These observations were confirmed by statistical and post-hoc analysis.

[Insert Figure 5 about here]

Discussion

Previous studies on reversal learning have used fMRI [25; 36; 45; 83; 92; 121] or EEG [105]. fNIRS has better temporal resolution than fMRI [30] and better spatial resolution than EEG

[119]. In this pilot study, we investigated the cortical hemodynamic response during a reversal-learning task using fNIRS.

We observed distinct neural substrates for reversal learning driven by reward and punishment using comparable magnitude feedback, as proposed by Xue et al., 2013 [121]. The l-DLPFC and l- IFG were involved in the reversal process when receiving unexpected positive feedback (reward condition). Unexpected negative feedback (punishment condition) led to significant hemodynamic changes for the OFC but only a tendency towards significancy was observed for the r-DLPFC and r-IFG.

Behavioral data from our pilot study show that all subjects understood the rules, as they succeeded in accumulating a positive number of points. Under the reward condition, participants made more errors and were faster to respond (especially after the reversal trials) than under the punishment condition. For the HbO parameter, we observed (1) longer and higher significant activation for the l-DLPFC for $t_{10\text{-}15}$ under the reward condition than under the punishment condition, (2) significantly greater activation in the l-DLPFC than r-DLPFC under the reward condition, (3) greater and prolonged activation in the OFC under the punishment condition, (4) significantly faster and a tendency for prolonged and greater activation of the r-DLPFC and greater and faster involvement of the r-IFG under punishment conditions. Finally, we observed significantly faster involvement of the right than left IFG for $t_{00\text{-}10}$ in terms of the HbR parameter under the punishment condition.

Contribution of the dorsolateral prefrontal cortex (DLPFC) in the reversal-learning task

We observed neurovascular coupling over the l-DLPFC under the reward condition during the RLT. However, our results also suggest the possible involvement of the r-DLPFC under punishment conditions.

Previously, neuroimaging studies on RLT reported involvement of the DLPFC either bilaterally [20; 25; 115] or only for the left [41] or right hemisphere [83; 121]. Concerning reinforcement learning, Xue et al. first revealed distinct mechanisms underlying learning from positive and negative feedback (i.e., reward and punishment, respectively) [121]. They observed rightward asymmetry (right lateral OFC and DLPFC) in punishment processing under punishment conditions, but no significant activation was observed for the reward condition, probably due to the small reward used in their study. Xue et al. [121], used various types of feedback (electric shocks and points later converted to dollars at a ratio of 25:1) associated with an incomparable level of magnitude between the two conditions. As suggested by Xue et al. [121], we used comparable feedback (gain or loss of money), with an equal level of magnitude (+100$ vs. -100$), in this pilot study to investigate local asymmetries in the

treatment of information to RLT specific to each condition (reward and punishment). Numerous studies have suggested that the DLPFC may be involved in various cognitive tasks: working memory [13], reasoning [90], changes in attention [37], and control [47]. Here, we observed lateralization of its activation according to the condition: l-DPLFC for the reward condition and r-DLPFC for the punishment condition. These observations support involvement of the DLPFC in the detection of switching [25; 92; 121], the updating of response-outcome relationships, and flexible behavior [45; 84], with lateralization according to the condition.

Contribution of the inferior frontal gyrus (IFG) in the reversal-learning task

The IFG is involved in inhibitory control [8; 34; 44; 65; 67; 68; 97; 109]. The IFG has also been shown to be involved in the RLT [47; 14], mostly in the right hemisphere [25*;* 45*;* 115]. Consistent with these previous results, we observed significant neurovascular coupling for the l-IFG under the reward condition, whereas greater and faster involvement of the r-IFG was observed under the punishment condition. These results support the role of the IFG in inhibiting a well-learned association [5; 45; 97] and also suggests lateralization according to the condition.

Contribution of the orbitofrontal cortex (OFC) in the reversal-learning task

We observed significant neurovascular coupling for the OFC under the punishment condition. O'Doherty et al., [83] also measured an increase in lateral OFC activity following the subjects' receipt of punishment and deactivation following reward. Although our mapping did not allow differentiation of the subregions of the OFC, we confirm involvement of the OFC under punishment conditions. The OFC is involved in maintaining the current and expected motivational value of stimuli [85; 117] and motivation-related processes [96; 107]. More precisely, the role of the OFC in reversal is to store the feedback association during reversal learning [17; 66; 78]. Numerous studies have argued that the right hemisphere plays a dominant role in experiencing unpleasant feelings, whereas the left hemisphere is essential for pleasant feelings [12; 32; 35; 86] and positive affects [11; 33; 52; 54]. Moreover, such lateralization has also been observed during reversal processing for the OFC, lateralized to the r-OFC, whereas reward has been found to be associated with the left hemisphere [105; 120]. These results support our hypothesis of lateralization for the DLPFC, IFG and, perhaps, the OFC. According to previous studies, the OFC provides information concerning the value of the stimulus to the DLPFC [110], which could then be used to select appropriate goals. The relationship is bidirectional [107]. Indeed, the involvement of the l-DLPFC (l-DLPFC under the reward and r-DLPFC under the punishment condition) appears to facilitate the updating of

response-outcome relationships and flexible behavior. Thus, the DLPFC may modulate value information stored in the OFC to be congruent with the current association [51; 107].

Overall cortical network activity in the reversal-learning task

The involvement of the IFG (l-IFG for the reward condition and r-IFG for the punishment condition) was linked to the inhibition of a well-learned association and implementation of behavioral rules or strategies. It is likely that such inhibition co-exists with other cognitive functions required by the RLT task (e.g., updating, shifting), making it difficult to establish which structures are involved in updating and inhibition processes. In 2007, Dosenbach et al. was able to distinguish between two strongly inter-connected sub-networks that function in parallel [39]. One involved the DLFPC, which is associated with top-down attentional control [39], maintaining goals, and updating information [116], whereas the other involved the anterior cingulate cortex (ACC), which is involved in detecting conflicting responses and monitoring performance [9; 23; 81; 103]. In our pilot study, involvement of the l-DLPFC is consistent with the involvement of this first sub-network. However, involvement of the ACC cannot be robustly investigated by fNIRS due to the low sensitivity of fNIRS for deep brain tissues.

The causality between the two sub-networks is not clear [40; 82; 106; 108] but studies have also suggested involvement of the subthalamic nucleus (STN), which is linked to the motor inhibition loop [6; 99]. The motor inhibition loop involves, among other elements, a "longer" indirect pathway (DLPFC-caudate-IFG-supplementary motor area -STN-motor area) related to the implementation of proactive modulation [7; 10; 18; 31; 113; 114]. In addition, the involvement of the IFG in our pilot study is suggestive of its participation in this inhibitory loop [63; 109]. Consistent with this hypothesis, the results concerning the slope coefficient suggest possible longer and greater activation of the DLPFC, which may then recruit the IFG, which would participate in the initiation of a motor inhibition loop. Based on recent studies, the connection between the IFG and DLPFC can be bidirectional. First, it is possible that the DLPFC recruits the IFG via an excitatory connection to initiate the longer indirect inhibitory pathway. Second, the IFG might then inhibit DLPFC activity during reappraisal once the strategy process is updated [9; 79]. This second activation of the IFG could explain the second peak observed for the r-IFG under the punishment condition and that of the l-IFG under the reward condition.

In summary, in this pilot study, we used fNIRS to investigate the differential impact of reversal: with involvement of the left hemisphere under the reward-guided condition, characterized by early l-DLPFC activation followed by involvement of the l-IFG and OFC,

and right hemisphere involvement (r-DLPFC and r-IFGC) under the punishment condition. The pilot study suggests feedback-specific processing during RLTs.

Limitations of the study

As for fMRI, we generally assume that there is a linear and time-invariant relationship between events of a specific task and the impulse response of the neurovascular system, called the hemodynamic response function (HRF). In common practice, it is generally accepted to adopt an a priori fixed shape for the HRF [73; 102]. For example, the canonical HRF is a standard choice, leaving only magnitude coefficients as free parameters to estimate. Such parametrization has the advantage of offering optimal statistical power and interpretability if the HRF shape is specified correctly. However, according to both the fMRI and fNIRS literature, there are temporal differences between Δ[HbO] and Δ[HbR] [56] and there is evidence that the shape of the HRF varies between subjects, brain regions, and tasks [3; 50; 53; 61; 74; 77]. Additionally, the HRF may also be altered in certain pathological conditions [75; 88; 101; 112]. For these reasons, fixed HRFs may not be an accurate a priori to model the dynamics of Δ[HbO] and Δ[HbR] in fNIRS. In addition, both 'GLM' and 'averaging' analysis procedures generate similar response morphologies and amplitude estimates [1;72]. Thus, to avoid incorporating a priori knowledge about the expected shape of the evoked hemodynamic response, we did not use GLM method.

Finally, it is worth mentioning that the influence of systemic blood-flow changes on hemodynamic signals is a potential problem for fNIRS data, although it may be more pronounced when recordings are made on freely moving people. fNIRS data consist of a combination of components arising from neuronal activity and those of systemic origin, which can lead to false positives and/or false negatives in the statistical inference of functional activity. Tachtsidis and Scholkmann [111] discussed this issue and how to address it in a recent review. The authors advised including approaches based on the various dynamic changes of both HbO and HbR to reduce confounding artefacts (as we performed) and allow better physiological interpretation of the functional experimental results.

The population of patients with DBS implants is often heterogenous, providing small subgroups for studies. We wished to assess whether the protocol works on groups with heterogeneous profiles (sex, age) and thus we did not restrict the population for this pilot study. One of the limitations of the pilot study was the small number of participants. Indeed, the present study was a preliminary analysis conducted before considering participants who had undergone DBS. Future studies with a larger sample size should be performed to validate our findings. Finally, we chose to use "hypothetical" money to have the same magnitude

between the two conditions and to avoid inducing explicit cognition and behavior biases between subjects due to the heterogeneity of the population.

Conclusion

In this pilot study, we wished to investigate and complement previous results concerning the brain regions involved in reversal learning.

The current study extends our understanding of the neural mechanisms involved in reversal learning and suggests feedback-specific processing during RLTs, with involvement of the left hemisphere under the reward-guided condition and the right hemisphere under the punishment condition. However, caution should be exercised when extrapolating results across studies.

The present work provides insights into the neurobiological mechanisms underlying behavior and brain (dys)function. In addition, the study is consistent with the concept of combining neurobiology, brain imaging modalities, and behavioral analysis to explore the spatial and temporal response of neural/vascular systems for subsequent use of the results in future clinical applications.

These findings, associated with future studies on neurological diseases characterized by a lack of flexibility (e.g., OCD or PD with DBS), will improve our comprehension of cortical neuronal and hemodynamic mechanisms as well as spatial involvement in such tasks. This approach demonstrates the possibility of using high-density optical imaging tools to study cognitive tasks. It can provide insights into neurophysiological mechanisms and facilitate the translation of functional optical imaging into clinical applications, such as in OCD or PD.

Competing interests

MM and FW have consulting activities for Seenel Imaging.

Funding

This work was funded by the ANR Physiobs and Région Hauts-de-France.

References

[1] Aarabi A, Osharina V, Wallois F. Effect of confounding variables on hemodynamic response function estimation using averaging and deconvolution analysis: An event-related NIRS study. NeuroImage 2017;155:25–49.

[2] Aasted CM, Yücel MA, Cooper RJ, Dubb J, Tsuzuki D, Becerra L, et al. Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. Neurophotonics 2015;2:020801.

[3] Aguirre GK, Zarahn E, D'esposito M. The variability of human, BOLD hemodynamic responses. Neuroimage 1998;8:360–9.

[4] Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. Impairment of social and moral behavior related to early damage in human prefrontal cortex. Nat Neurosci 1999;2:1032–7.

[5] Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex. Trends Cogn Sci (Regul Ed) 2004;8:170–7.

[6] Aron AR, Poldrack RA. Cortical and Subcortical Contributions to Stop Signal Response Inhibition: Role of the Subthalamic Nucleus. J Neurosci 2006;26:2424–33.

[7] Aron AR. From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. Biol Psychiatry 2011;69:e55-68.

[8] Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex: one decade on. Trends Cogn Sci (Regul Ed) 2014;18:177–85.

[9] Banich MT. The Stroop Effect Occurs at Multiple Points Along a Cascade of Control: Evidence From Cognitive Neuroscience Approaches. Front Psychol 2019;10.

[10] Bari A, Robbins TW. Inhibition and impulsivity: behavioral and neural basis of response control. Prog Neurobiol 2013;108:44–79.

[11] Baxter LR, Schwartz JM, Phelps ME, Mazziotta JC, Guze BH, Selin CE, et al. Reduction of prefrontal cortex glucose metabolism common to three types of depression. Arch Gen Psychiatry 1989;46:243–50.

[12] Bechara A, Damasio AR. The somatic marker hypothesis: A neural theory of economic decision. Games and Economic Behavior 2005;52:336–72.

[13] Belger A, Puce A, Krystal JH, Gore JC, Goldman-Rakic P, McCarthy G. Dissociation of mnemonic and perceptual processes during spatial and nonspatial working memory using fMRI. Hum Brain Mapp 1998;6:14–32.

[14] Budhani S, Marsh AA, Pine DS, Blair RJR. Neural correlates of response reversal: considering acquisition. Neuroimage 2007;34:1754–65.

[15] Butter CM, Mishkin M, Rosvold HE. Conditioning and extinction of a food-rewarded response after selective ablations of frontal cortex in rhesus monkeys. Experimental Neurology 1963;7:65–75. https://doi.org/10.1016/0014-4886(63)90094-3.

[16] Buxton RB. The Elusive Initial Dip. NeuroImage 2001;13:953–8.

[17] Cai X, Padoa-Schioppa C. Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. Neuron 2014;81:1140–51.

[18] Cai Y, Li S, Liu J, Li D, Feng Z, Wang Q, et al. The Role of the Frontal and Parietal Cortex in Proactive and Reactive Inhibitory Control: A Transcranial Direct Current Stimulation Study. J Cogn Neurosci 2016;28:177–86.

[19] Chabardes S, Krack P, Piallat B, Bougerol T, Seigneuret E, Yelnik J, et al. Deep brain stimulation of the subthalamic nucleus in obsessive–compulsives disorders: long-term follow-up of an open, prospective, observational cohort. J Neurol Neurosurg Psychiatry 2020;91:1349–56.

[20] Chamberlain SR, Menzies L, Hampshire A, Suckling J, Fineberg NA, del Campo N, et al. Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. Science 2008;321:421–2.

[21] Chudasama Y, Daniels TE, Gorrin DP, Rhodes SEV, Rudebeck PH, Murray EA. The Role of the Anterior Cingulate Cortex in Choices based on Reward Value and Reward Contingency. Cerebral Cortex 2013;23:2884–98.

[22] Clark L, Manes F, Antoun N, Sahakian BJ, Robbins TW. The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. Neuropsychologia 2003;41:1474–83.

[23] Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS. Multi-task connectivity reveals flexible hubs for adaptive task control. Nat Neurosci 2013;16:1348–55.

[24] Cools R, Barker RA, Sahakian BJ, Robbins TW. Mechanisms of cognitive set flexibility in Parkinson's disease. Brain 2001;124:2503–12.

[25] Cools R, Clark L, Owen AM, Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. J Neurosci 2002;22:4563–7.

[26] Cools R. Dopaminergic modulation of cognitive function-implications for l-DOPA treatment in Parkinson's disease. Neuroscience & Biobehavioral Reviews 2006;30:1–23.

[27] Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M. Striatal Dopamine Predicts Outcome-Specific Reversal Learning and Its Sensitivity to Dopaminergic Drug Administration. J Neurosci 2009;29:1538–43.

[28] Cooper RJ, Hebden JC, O'Reilly H, Mitra S, Michell AW, Everdell NL, et al. Transient haemodynamic events in neurologically compromised infants: a simultaneous EEG and diffuse optical imaging study. Neuroimage 2011;55:1610–6.

[29] Costantini M, Di Vacri A, Chiarelli AM, Ferri F, Luca Romani G, Merla A. Studying social cognition using near-infrared spectroscopy: the case of social Simon effect. J Biomed Opt 2013;18:25005.

[30] Cui X, Bray S, Bryant DM, Glover GH, Reiss AL. A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. Neuroimage 2011;54:2808–21.

[31] Cunillera T, Fuentemilla L, Brignani D, Cucurell D, Miniussi C. A Simultaneous Modulation of Reactive and Proactive Inhibition Processes by Anodal tDCS on the Right Inferior Frontal Cortex. PLOS ONE 2014;9:e113537.

[32] Davidson RJ, Irwin W. The functional neuroanatomy of emotion and affective style. Trends in Cognitive Sciences 1999;3:11–21.

[33] Davidson RJ, Henriques J. Regional brain function in sadness and depression. The neuropsychology of emotion, New York, NY, US: Oxford University Press; 2000, p. 269–97.

[34] de Zubicaray GI, Andrew C, Zelaya FO, Williams SC, Dumanoir C. Motor response suppression and the prepotent tendency to respond: a parametric fMRI study. Neuropsychologia 2000;38:1280–91.

[35] Deglin VL, Kinsbourne M. Divergent Thinking Styles of the Hemispheres: How Syllogisms Are Solved during Transitory Hemisphere Suppression. Brain and Cognition 1996;31:285–307.

[36] Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA. Tracking the hemodynamic responses to reward and punishment in the striatum. J Neurophysiol 2000;84:3072–7.

[37] Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. Nature 1996;380:69–72.

[38] Dickstein DP, Finger EC, Brotman MA, Rich BA, Pine DS, Blair JR, et al. Impaired probabilistic reversal learning in youths with mood and anxiety disorders. Psychol Med 2010;40:1089–100.

[39] Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, et al. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 2007;104:11073–8.

[40] Duann J-R, Ide JS, Luo X, Li CR. Functional Connectivity Delineates Distinct Roles of the Inferior Frontal Cortex and Presupplementary Motor Area in Stop Signal Inhibition. J Neurosci 2009;29:10171–9.

[41] Fellows LK, Farah MJ. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain 2003;126:1830–7.

[42] Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 2012;63:921–35.

[43] Frank MJ, Seeberger LC, O'Reilly RC. By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. Science 2004;306:1940–3.

[44] Garavan H, Ross TJ, Stein EA. Right hemispheric dominance of inhibitory control: an event-related functional MRI study. Proc Natl Acad Sci USA 1999;96:8301–6.

[45] Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA. Neural Components Underlying Behavioral Flexibility in Human Reversal Learning. Cereb Cortex 2010;20:1843–52.

[46] Gruner P, Pittenger C. Cognitive inflexibility in Obsessive-Compulsive Disorder. Neuroscience 2017;345:243–55.

[47] Hampshire A, Owen AM. Fractionating attentional control using event-related fMRI. Cereb Cortex 2006;16:1679–89.

[48] Hampshire A, Chaudhry AM, Owen AM, Roberts AC. Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. NeuroImage 2012;59:4102–12.

[49] Hampton AN, O'Doherty JP. Decoding the neural substrates of reward-related decision making with functional MRI. Proc Natl Acad Sci U S A 2007;104:1377–82.

[50] Handwerker DA, Gonzalez-Castillo J, D'Esposito M, Bandettini PA. The continuing challenge of understanding and modeling hemodynamic variation in fMRI. Neuroimage 2012;62:1017–23.

[51] Hare TA, Camerer CF, Rangel A. Self-control in decision-making involves modulation of the vmPFC valuation system. Science 2009;324:646–8.

[52] Harmon-Jones E, Allen JJ. Behavioral activation sensitivity and resting frontal EEG asymmetry: covariation of putative indicators related to risk for mood disorders. J Abnorm Psychol 1997;106:159–63.

[53] Heinzel S, Metzger FG, Ehlis A-C, Korell R, Alboji A, Haeussinger FB, et al. Aging-related cortical reorganization of verbal fluency processing: a functional near-infrared spectroscopy study. Neurobiol Aging 2013;34:439–50.

[54] Herrington JD, Mohanty A, Koven NS, Fisher JE, Stewart JL, Banich MT, et al. Emotion-modulated performance and activity in left dorsolateral prefrontal cortex. Emotion 2005;5:200–7.

[55] Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, et al. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. J Cogn Neurosci 2004;16:463–78.

[56] Huppert TJ, Hoge RD, Diamond SG, Franceschini MA, Boas DA. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. Neuroimage 2006;29:368–82.

[57] Huppert TJ, Diamond SG, Franceschini MA, Boas DA. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. Appl Opt 2009;48:D280-298.

[58] Izquierdo A, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J Neurosci 2004;24:7540–8.

[59] Izquierdo A, Jentsch JD. Reversal learning as a measure of impulsive and compulsive behavior in addictions. Psychopharmacology (Berl) 2012;219:607–20.

[60] Izquierdo A, Brigman JL, Radke AK, Rudebeck PH, Holmes A. The neural basis of reversal learning: An updated perspective. Neuroscience 2017;345:12–26.

[61] Jasdzewski G, Strangman G, Wagner J, Kwong KK, Poldrack RA, Boas DA. Differences in the hemodynamic response to event-related motor and visual paradigms as measured by near-infrared spectroscopy. NeuroImage 2003;20:479–88.

[62] Jöbsis FF. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. Science 1977;198:1264–7.

[63] Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. Neuroscience 2006;139:181–93.

[64] Kawai T, Yamada H, Sato N, Takada M, Matsumoto M. Roles of the Lateral Habenula and Anterior Cingulate Cortex in Negative Outcome Monitoring and Behavioral Adjustment in Nonhuman Primates. Neuron 2015;88:792–804.

[65] Kawashima R, Satoh K, Itoh H, Ono S, Furumoto S, Gotoh R, et al. Functional anatomy of GO/NO-GO discrimination and response selection--a PET study in man. Brain Res 1996;728:79–89.

[66] Keiflin R, Reese RM, Woods CA, Janak PH. The Orbitofrontal Cortex as Part of a Hierarchical Neural System Mediating Choice between Two Good Options. J Neurosci 2013;33:15989–98.

[67] Konishi S, Nakajima K, Uchida I, Sekihara K, Miyashita Y. No-go dominant brain activity in human inferior prefrontal cortex revealed by functional magnetic resonance imaging. Eur J Neurosci 1998;10:1209–13.

[68] Konishi S, Nakajima K, Uchida I, Kikyo H, Kameyama M, Miyashita Y. Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. Brain 1999;122 (Pt 5):981–91.

[69] Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. Gainesville, Fla.: NIMH, Center for the Study of Emotion & Attention; 2008.

[70] Lee EC, Whitehead AL, Jacques RM, Julious SA. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? BMC Med Res Methodol 2014;14:41.

[71] León-Carrión J, León-Domínguez U. Functional Near-Infrared Spectroscopy (fNIRS): Principles and Neuroscientific Applications. Neuroimaging - Methods 2012.

[72] Luke R, Larson ED, Shader MJ, Innes-Brown H, Yper LV, Lee AKC, et al. Analysis methods for measuring passive auditory fNIRS responses generated by a block-design paradigm. NPh 2021;8:025008.

[73] Machado A, Lina JM, Tremblay J, Lassonde M, Nguyen DK, Lesage F, et al. Detection of hemodynamic responses to epileptic activity using simultaneous Electro-EncephaloGraphy (EEG)/Near Infra Red Spectroscopy (NIRS) acquisitions. NeuroImage 2011;56:114–25.

[74] Mahmoudzadeh M, Dehaene-Lambertz G, Fournier M, Kongolo G, Goudjil S, Dubois J, et al. Syllabic discrimination in premature human infants prior to complete formation of cortical layers. PNAS 2013;110:4846–51.

[75] Mahmoudzadeh M, Dehaene-Lambertz G, Kongolo G, Fournier M, Goudjil S, Wallois F. Consequence of intraventricular hemorrhage on neurovascular coupling evoked by speech syllables in preterm neonates. Dev Cogn Neurosci 2018;30:60–9.

[76] Mandrick K, Derosiere G, Dray G, Coulon D, Micallef J-P, Perrey S. Utilizing slope method as an alternative data analysis for functional near-infrared spectroscopy-derived cerebral hemodynamic responses. International Journal of Industrial Ergonomics 2013;43:335–41.

[77] Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 2000;11:735–59.

[78] Moorman DE, Aston-Jones G. Orbitofrontal Cortical Neurons Encode Expectation-Driven Initiation of Reward-Seeking. J Neurosci 2014;34:10234–46.

[79] Morawetz C, Bode S, Baudewig J, Kirilina E, Heekeren HR. Changes in Effective Connectivity Between Dorsal and Ventral Prefrontal Regions Moderate Emotion Regulation. Cereb Cortex 2016;26:1923–37.

[80] Nagahama Y, Okada T, Katsumi Y, Hayashi T, Yamauchi H, Oyanagi C, et al. Dissociable mechanisms of attentional control within the human prefrontal cortex. Cereb Cortex 2001;11:85–92.

[81] Nelson JK, Reuter-Lorenz PA, Sylvester C-YC, Jonides J, Smith EE. Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. PNAS 2003;100:11171–5.

[82] Neubert F-X, Mars RB, Buch ER, Olivier E, Rushworth MFS. Cortical and subcortical interactions during action reprogramming and their related white matter pathways. Proc Natl Acad Sci U S A 2010;107:13240–5.

[83] O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci 2001;4:95–102.

[84] O'Doherty J, Critchley H, Deichmann R, Dolan RJ. Dissociating Valence of Outcome from Behavioral Control in Human Orbital and Ventral Prefrontal Cortices. J Neurosci 2003;23:7931–9.

[85] O'Doherty JP, Dolan RJ. The role of human orbitofrontal cortex in reward prediction and behavioral choice: insights from neuroimaging. Oxford University Press; 2006

[86] Overskeid G. The Slave of the Passions: Experiencing Problems and Selecting Solutions. Review of General Psychology 2000;4:284–309.

[87] Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, et al. Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning. Neuron 2012;76:998–1009.

[88] Pellegrino G, Machado A, von Ellenrieder N, Watanabe S, Hall JA, Lina J-M, et al. Hemodynamic Response to Interictal Epileptiform Discharges Addressed by Personalized EEG-fNIRS Recordings. Front Neurosci 2016;10:102.

[89] Perrey S, Thedon T, Rupp T. NIRS in ergonomics: Its application in industry for promotion of health and human performance at work. International Journal of Industrial Ergonomics 2010;40:185–9.

[90] Prabhakaran V, Smith JA, Desmond JE, Glover GH, Gabrieli JD. Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. Cogn Psychol 1997;33:43–63.

[91] Reber J, Tranel D. Sex differences in the functional lateralization of emotion and decision-making in the human brain. J Neurosci Res 2017;95:270–8.

[92] Remijnse PL, Nielen MMA, Uylings HBM, Veltman DJ. Neural correlates of a reversal learning task with an affectively neutral baseline: an event-related fMRI study. Neuroimage 2005;26:609–18.

[93] Remijnse PL, Heuvel OA van den, Nielen MMA, Vriend C, Hendriks G-J, Hoogendijk WJG, et al. Cognitive Inflexibility in Obsessive-Compulsive Disorder and Major Depression Is Associated with Distinct Neural Correlates. PLOS ONE 2013;8:e59600.

[94] Robinson OJ, Frank MJ, Sahakian BJ, Cools R. Dissociable responses to punishment in distinct striatal regions during reversal learning. NeuroImage 2010;51:1459–67.

[95] Roche-Labarbe N, Zaaimi B, Berquin P, Nehlig A, Grebe R, Wallois F. NIRS-measured oxy- and deoxyhemoglobin changes associated with EEG spike-and-wave discharges in children. Epilepsia 2008;49:1871–80.

[96] Rothkirch M, Schmack K, Schlagenhauf F, Sterzer P. Implicit motivational value and salience are processed in distinct areas of orbitofrontal cortex. Neuroimage 2012;62:1717–25.

[97] Rygula R, Walker SC, Clarke HF, Robbins TW, Roberts AC. Differential Contributions of the Primate Ventrolateral Prefrontal and Orbitofrontal Cortex to Serial Reversal Learning. J Neurosci 2010;30:14552–9.

[98] Saver JL, Damasio AR. Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. Neuropsychologia 1991;29:1241–9.

[99] Schmidt R, Leventhal DK, Mallet N, Chen F, Berke JD. Canceling actions involves a race between basal ganglia pathways. Nat Neurosci 2013;16:1118–24.

[100] Scholkmann F, Spichtig S, Muehlemann T, Wolf M. How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. Physiol Meas 2010;31:649–62.

[101] Schroeter ML, Zysset S, Kruggel F, von Cramon DY. Age dependency of the hemodynamic response as measured by functional near-infrared spectroscopy. Neuroimage 2003;19:555–64.

[102] Schroeter ML, Bücheler MM, Müller K, Uludağ K, Obrig H, Lohmann G, et al. Towards a standard analysis for functional near-infrared imaging. NeuroImage 2004;21:283–90.

[103] Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, et al. Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. J Neurosci 2007;27:2349–56.

[104] Smielewski P, Kirkpatrick P, Minhas P, Pickard JD, Czosnyka M. Can cerebrovascular reactivity be measured with near-infrared spectroscopy? Stroke 1995;26:2285–92.

[105] Sobotka SS, Davidson RJ, Senulis JA. Anterior brain electrical asymmetries in response to reward and punishment. Electroencephalography and Clinical Neurophysiology 1992;83:236–47.

[106] Spechler PA, Chaarani B, Hudson KE, Potter A, Foxe JJ, Garavan H. Response inhibition and addiction medicine: from use to abstinence. Prog Brain Res 2016;223:143–64.

[107] Spielberg JM, Miller GA, Warren SL, Engels AS, Crocker LD, Banich MT, et al. A brain network instantiating approach and avoidance motivation. Psychophysiology 2012;49:1200–14.

[108] Swann NC, Cai W, Conner CR, Pieters TA, Claffey MP, George JS, et al. Roles for the pre-supplementary motor area and the right inferior frontal gyrus in stopping action: electrophysiological responses and functional and structural connectivity. Neuroimage 2012;59:2860–70.

[109] Swick D, Ashley V, Turken AU. Left inferior frontal gyrus is critical for response inhibition. BMC Neurosci 2008;9:102.

[110] Szatkowska I, Bogorodzki P, Wolak T, Marchewka A, Szeszkowski W. The effect of motivation on working memory: An fMRI and SEM study. Neurobiology of Learning and Memory 2008;90:475–8.

[111] Tachtsidis I, Scholkmann F. False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. Neurophoton 2016;3:031405.

[112] Tak S, Yoon SJ, Jang J, Yoo K, Jeong Y, Ye JC. Quantitative analysis of hemodynamic and metabolic changes in subcortical vascular dementia using simultaneous near-infrared spectroscopy and fMRI measurements. NeuroImage 2011;55:176–84.

[113] Tops M, Boksem MAS, Quirin M, IJzerman H, Koole SL. Internally directed cognition and mindfulness: an integrative perspective derived from predictive and reactive control systems theory. Front Psychol 2014;5.

[114] Verbruggen F, Logan GD. Automatic and Controlled Response Inhibition: Associative Learning in the Go/No-Go and Stop-Signal Paradigms. J Exp Psychol Gen 2008;137:649–72.
[115] Waegeman A, Declerck CH, Boone C, Seurinck R, Parizel PM. Individual differences in behavioral flexibility in a probabilistic reversal learning task: An fMRI study. Journal of Neuroscience, Psychology, and Economics 2014;7:203–18.
[116] Wager TD, Smith EE. Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci 2003;3:255–74.
[117] Wallis JD. Orbitofrontal cortex and its contribution to decision-making. Annu Rev Neurosci 2007;30:31–56.
[118] Wheeler EZ, Fellows LK. The human ventromedial frontal lobe is critical for learning from negative feedback. Brain 2008;131:1323–31.
[119] Wilcox T, Biondi M. fNIRS in the developmental sciences. Wiley Interdiscip Rev Cogn Sci 2015;6:263–83.
[120] Xue G, Lu Z, Levin IP, Weller JA, Li X, Bechara A. Functional Dissociations of Risk and Reward Processing in the Medial Prefrontal Cortex. Cereb Cortex 2009;19:1019–27.
[121] Xue G, Xue F, Droutman V, Lu Z-L, Bechara A, Read S. Common neural mechanisms underlying reversal learning by reward and punishment. PLoS ONE 2013;8:e82169.

Figure legends

Figure 1. Reversal-learning task. (A) The task was composed of four task blocks of 360 s and two training blocks of 180 s. **(B)** One task block lasted 360 s and was composed of 80 trials (40 positive and 40 negative). **(B-1)** Negative pair: antelope was the incorrect (-100 $) and cow the correct (+0 $) answer. **(B-2)** Positive pair: butterfly was the incorrect answer (+0 $), and duck was the correct answer (+100 $). The switch occurred after four to six correct answers (randomized) and the feedback was reversed. Thus, during the reversal phase, antelope became the correct answer (gain: 0 $) and cow the incorrect answer (gain: -100 $) for the negative pair **(B-3)** and duck became the incorrect answer (gain: 0 $) and butterfly the correct answer (gain: +100 $) for the positive pair **(B-4)**. We labelled each trial for each condition. During the acquisition phase, we distinguished between correct (ACR) and incorrect responses (AE). After 4 to 6 correct responses, a reversal occurred, unknown to the subject. During the reversal phase, we classified the responses as reversal errors (RE), the final reversal error (FRE) before the first correct response (RCR), and the reversal errors after the first CR and thus not linked to the switch (RENS). **(C)** For each 80 trials of one block, the trial consisted of a presentation of a pair of animal pictures (the stimuli), which was shown until the participant's response or a maximum of 3,000 ms, a short feedback period, and an interstimulus interval (ISI) of 1,000-1,500 ms (randomly determined).

Figure 2. Probe design used for the task of reversal learning (A) Optode positions: sources are in red and detectors in blue. The position of each optode is associated with a position of the international 10-10 EEG system. **(B)** Projection of the probe layout and meshing between sources and detectors: source 1 corresponds to point Fp1 and detector 2 to point Fp2. Each yellow line connecting a source and detector modeled one channel. **(C)** Sensitivity map associated with our probe design. **(D)** Identification of regions of interested identified using AtlasViewer. **(E)** The Medelopt® fNIRS setup used.

Figure 3. Averaged behavioral results. (A) The averaged accuracy and **(B)** reaction time for each repetition of pairs on the screen during the acquisition and reversal phases were plotted separately for the reward and punishment conditions. Each presentation, between two switches, consists of the average of eight participants' answers for the four recording blocks, thus at least 32 acquisitions. The perseverative errors (**C**) indicate the eight participants' average number of successive incorrect answers after the reversal for the four recordings

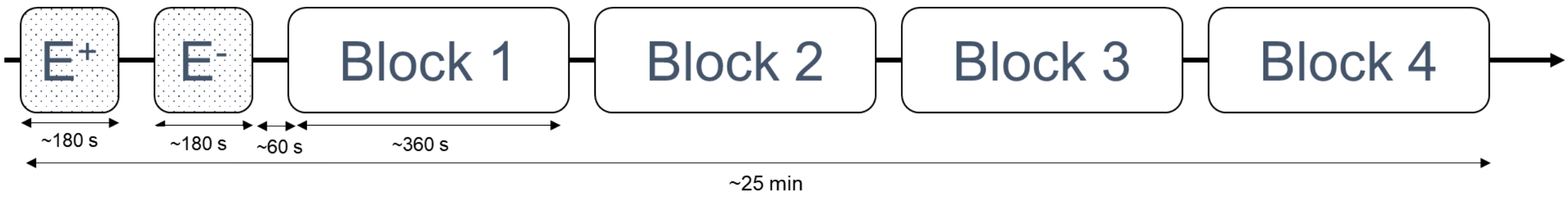
blocks; a lower score indicates that participants understood the reversal very quickly and applied it at the second repetition.

Figure 4. fNIRS averaged responses under the reward and punishment condition. (A) The right and left grand average of HbO (red) and HbR (blue) responses for both conditions are plotted for seven ROIs, which locations are presented over the right and left hemispheres. For clarity, the total-Hb (HbT) curve is not presented. Significant differences relative to baseline t_{-5-00} are indicated within their specific subperiods in red for HbO and blue for HbR. The contrast indicates the degree of significance (opaque: $p < 0.01$, transparent: $p < 0.05$). In the middle of the figure, only the projected HbO concentrations under the reward (top panel) and punishment conditions (bottom panel) for t_{05-15} are shown. Regions shaded in blue indicate under-activation, whereas those shaded in red indicate strong activation. The area of the shaded error bars around the curves indicates the standard deviation for the corresponding signals at each time-point. **(B) Hemodynamic activity of the left DLPFC,** comparing the average HbO and HbT response between the reward (POS) and punishment (NEG) conditions and the AUC for various subperiods. There was no significant difference for HbR. The t-test *p-values* are shown below for the reward condition relative to the punishment condition **(C) Hemodynamic activity of the right DLPFC,** comparing the HbO and HbT response between the reward (POS) and punishment (NEG) conditions. There was no significant difference for HbR. The t-test *p-values* are shown below for the reward condition relative to the punishment condition. **(D) Hemodynamic activity of the OFC,** comparing the average HbO and HbT response between the reward (POS) and punishment (NEG) conditions and AUC for various subperiods. There was no significant difference for HbR. The t-test *p-values* are shown below for the reward condition relative to the punishment condition. **(E) Hemodynamic activity of the l-IFG,** comparing the HbO response between the reward (POS) and punishment (NEG) conditions. There was no significant difference for HbR and HbT. The t-test *p-values* are shown below for the reward condition relative to the punishment condition. **(F) Hemodynamic activity of the r-IFG,** comparing the HbR response between the reward (POS) and punishment (NEG) conditions. There was no significant difference for HbO and HbT. The t-test *p-values* are shown below for the reward condition relative to the punishment condition.

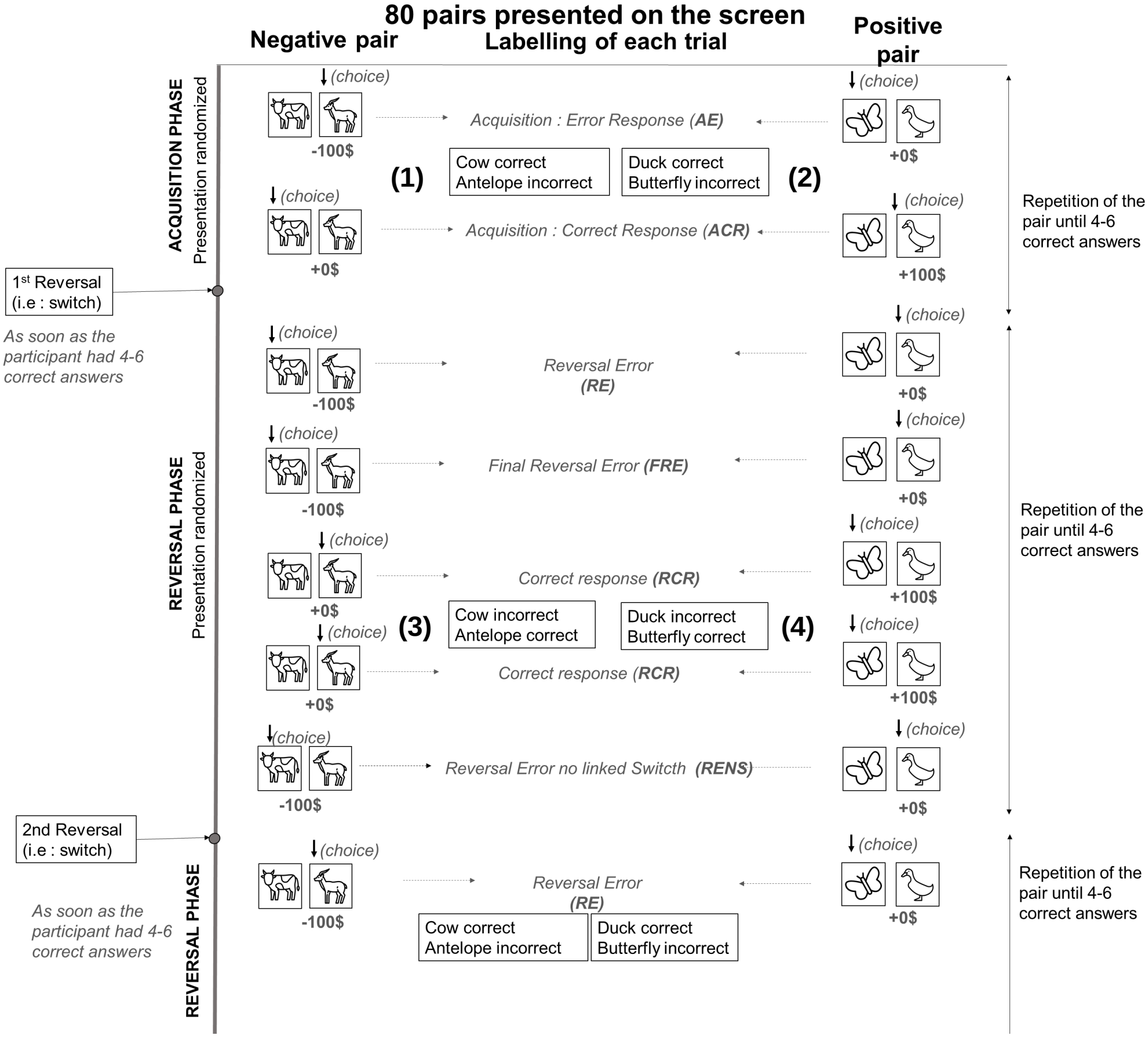
Figure 5. (A) Fisher LSD post hoc analysis for HbO for t_{05-15}. The figure is composed of three cross-tables, one for each parameter. Significant differences between two ROIs (line vs

columns) is symbolized by one or several asterix (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$), and *t* a tendency towards significance with a p-value < 0.09. Values are negative for right < left under the reward condition. Values are positive for reward > punishment when comparing the reward to punishment condition. **(B) Hemodynamic average difference**, in post-hoc analysis for HbO between the reward and punishment conditions. The difference is indicated by a red line with an *: * indicates a p-value < 0.05 and t a tendency towards significance, with a *p*-value < 0.09). The HbO concentration projection under the reward (top panel) and punishment conditions (bottom panel) for $t_{05\text{-}20}$ are shown. Regions shaded in blue indicate under-activation, whereas those shaded in red indicate strong activation. Area legend: 1: l-PFC, 2: l-DLPFC, 3: l-IFG, 4: OFC, 5: r-PFC, 6: r-DLPFC, 7: r-IFG. **(C) Summary of the fNIRS results.** This table summarizes all observations across several analyses: changes in amplitude and AUC relative to baseline activity under the reward (column 1) and punishment (column 2) conditions, box-plots comparing (column 3) the reward and punishment conditions for amplitude, AUC, and slope, and the post-hoc analyses (column 4).

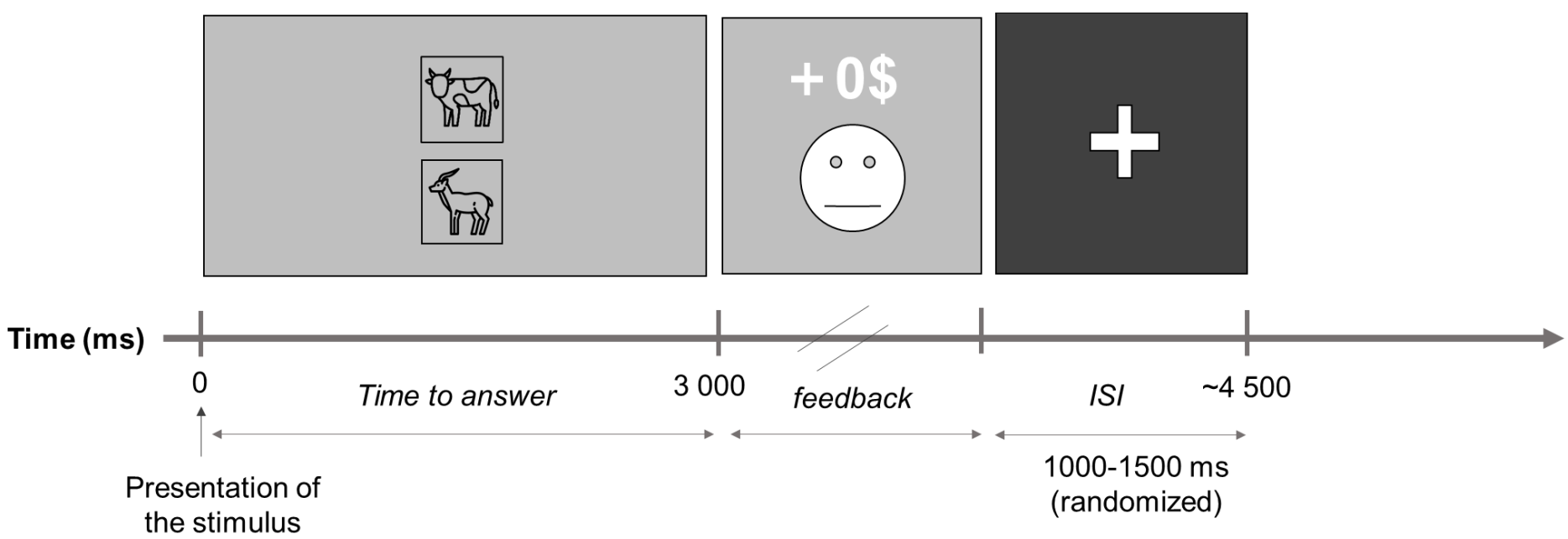
(A) Composition of the task

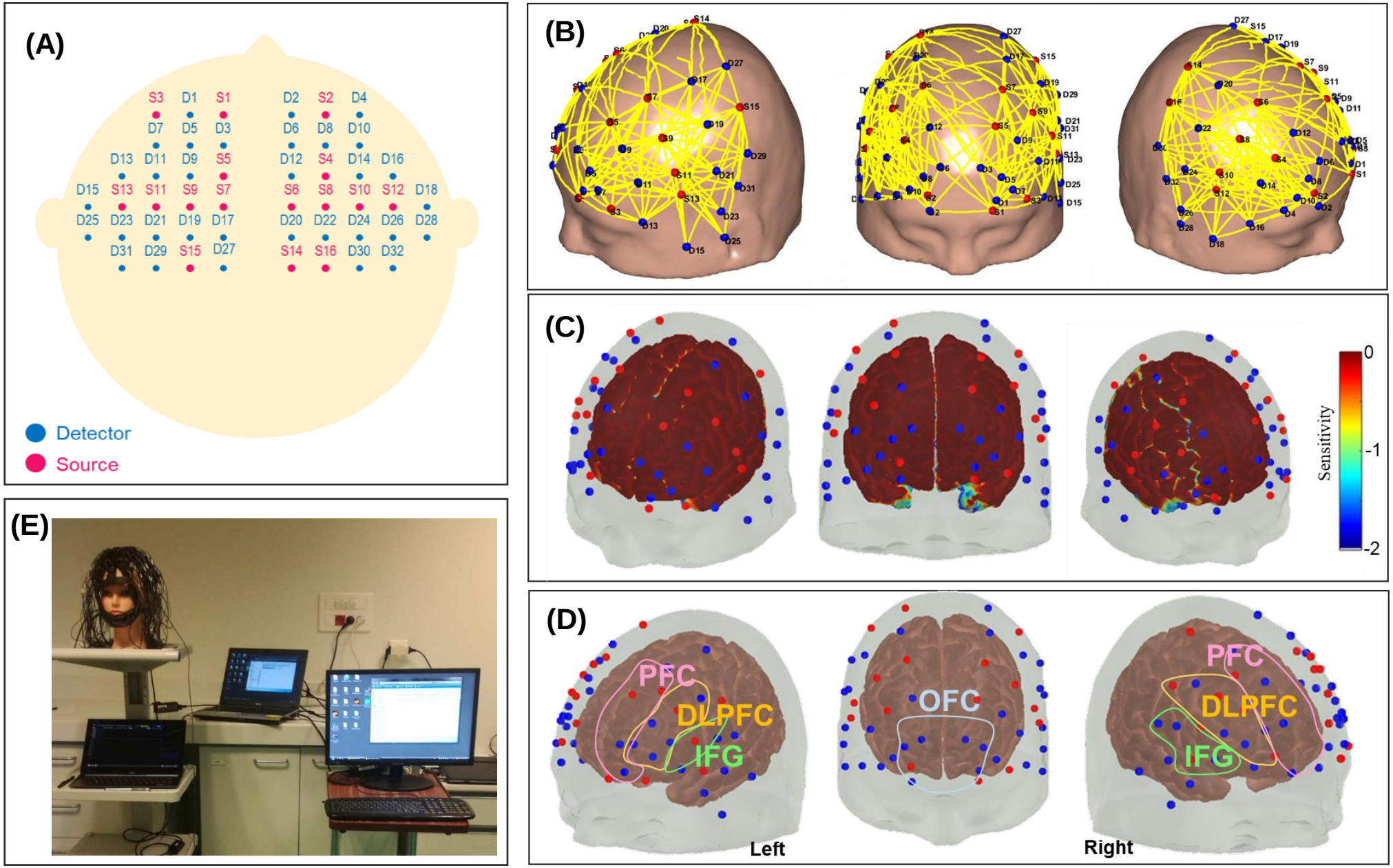


(B) Composition of one block



(C) Composition of one trial




(A)
S3 D1 S1
D2 S2 D4
D7 D5 D3
D6 D8 D10
D13 D11 D9 S5
D12 S4 D14 D16
D15 S13 S11 S9 S7
S6 S8 S10 S12 D18
D25 D23 D21 D19 D17
D20 D22 D24 D26 D28
D31 D29 S15 D27
S14 S16 D30 D32
Detector
Source
(B)
(C)
Sensitivity
0
-1
-2
(D)
PFC
DLPFC
IFG
OFC
Left
Right
(E)

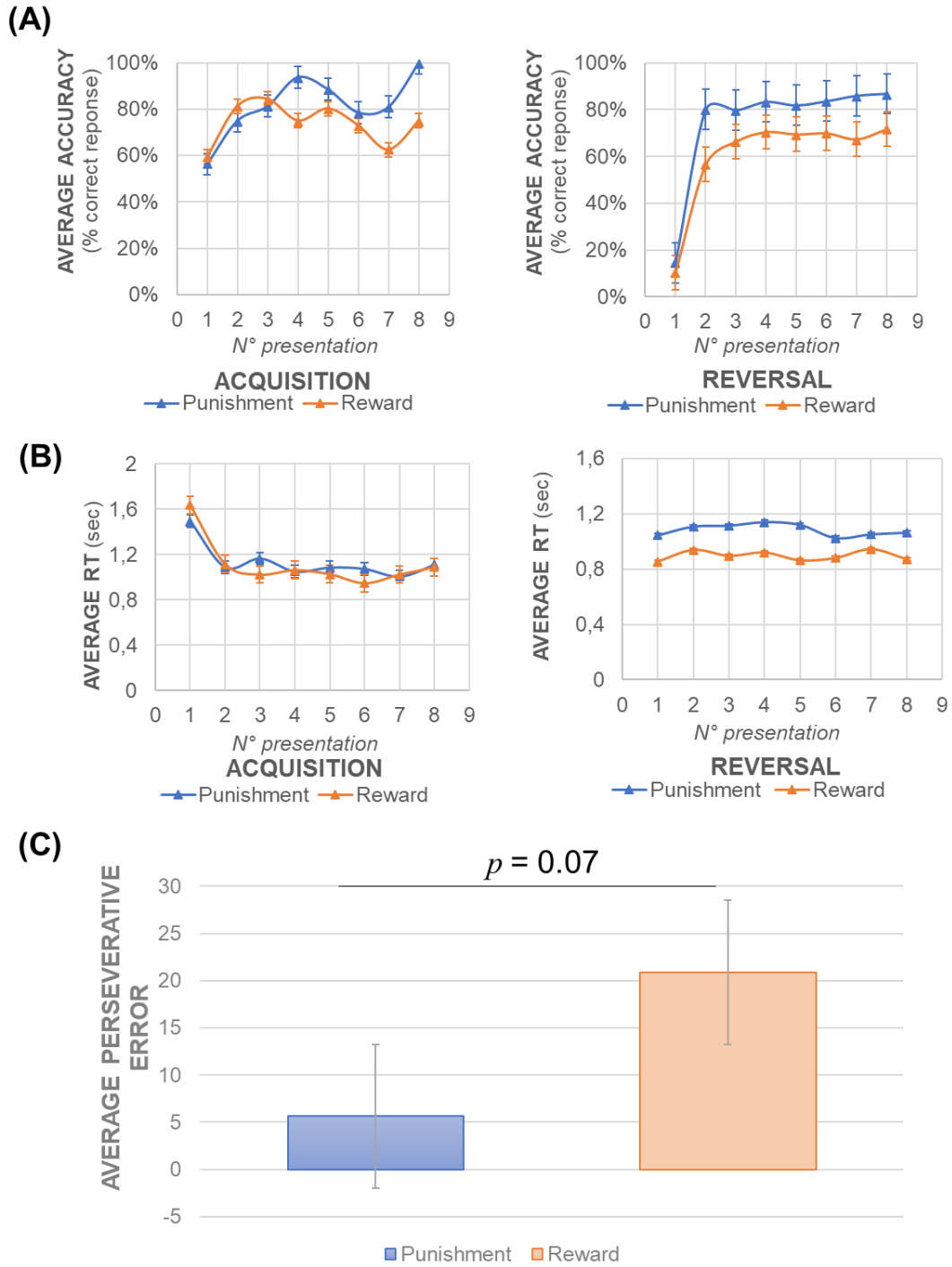
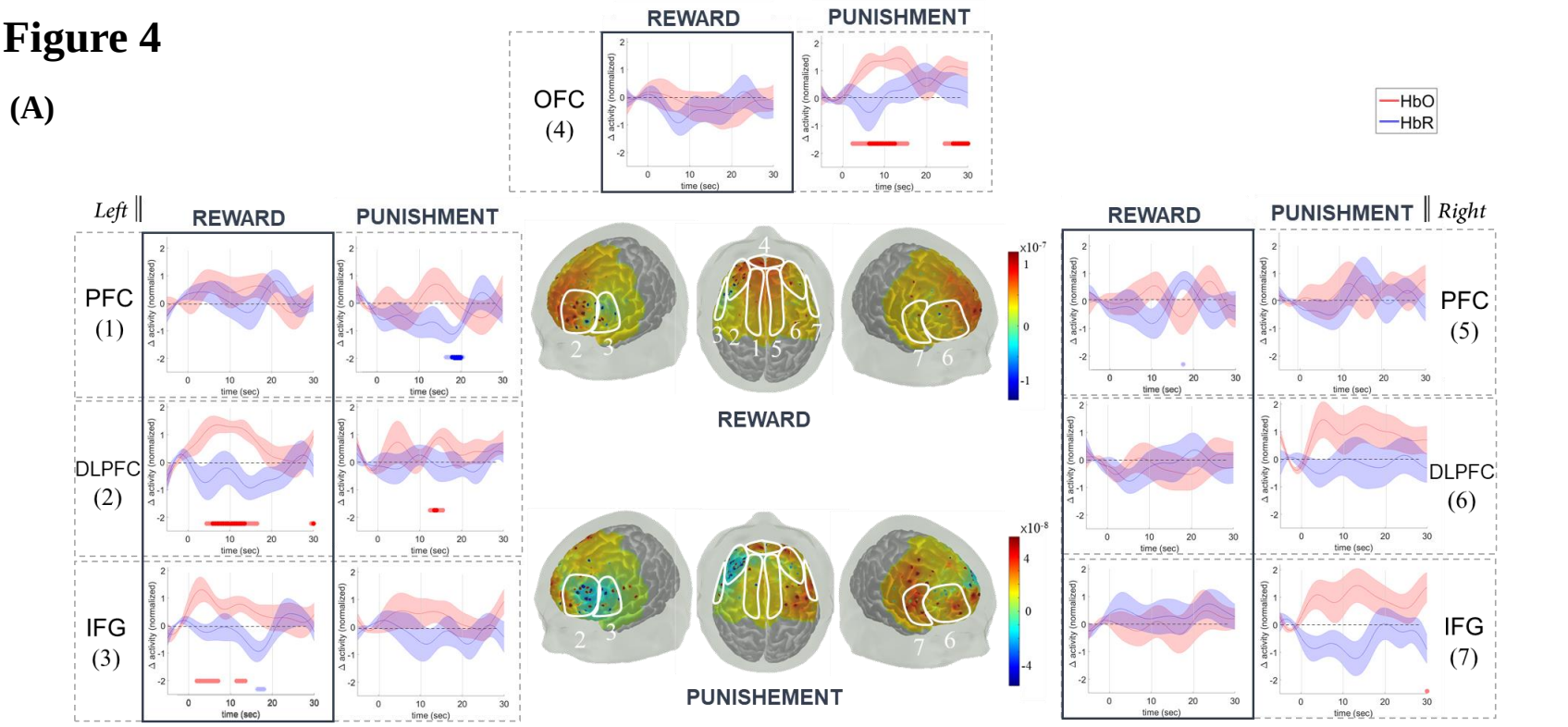
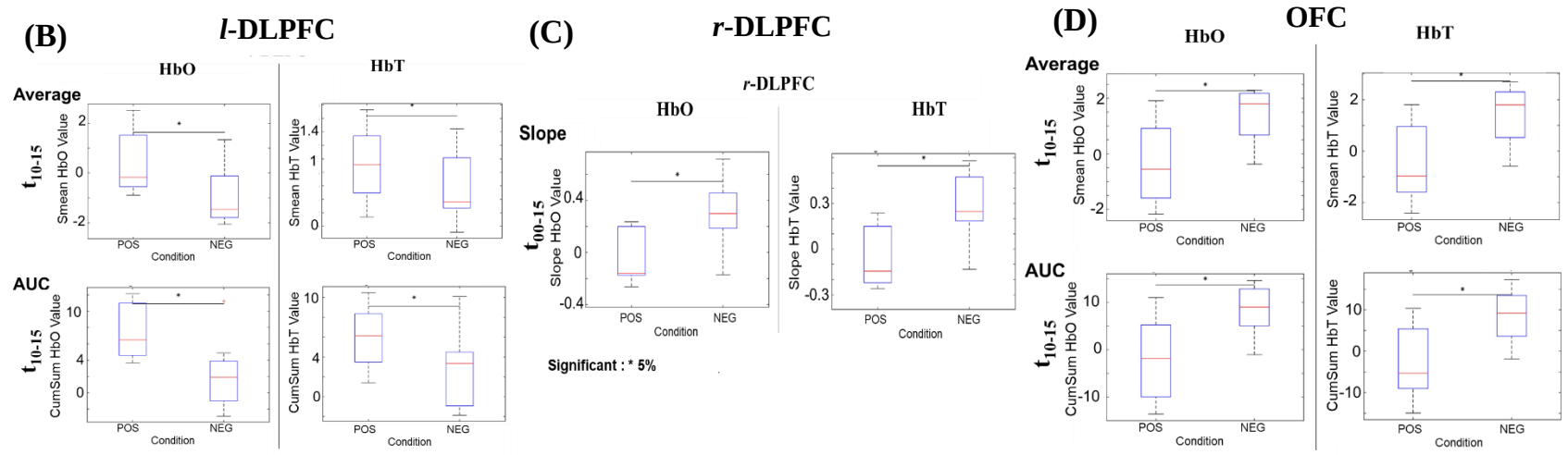

(A)
AVERAGE ACCURACY (% correct reponse)
100%
80%
60%
40%
20%
0%
0 1 2 3 4 5 6 7 8 9
N° presentation
ACQUISITION
Punishment
Reward
AVERAGE ACCURACY (% correct reponse)
100%
80%
60%
40%
20%
0%
0 1 2 3 4 5 6 7 8 9
N° presentation
REVERSAL
Punishment
Reward
(B)
AVERAGE RT (sec)
2
1,6
1,2
0,8
0,4
0
0 1 2 3 4 5 6 7 8 9
N° presentation
ACQUISITION
Punishment
Reward
AVERAGE RT (sec)
1,6
1,2
0,8
0,4
0
0 1 2 3 4 5 6 7 8 9
N° presentation
REVERSAL
Punishment
Reward
(C)
p = 0.07
AVERAGE PERSEVERATIVE ERROR
30
25
20
15
10
5
0
-5
Punishment
Reward

Figure 4

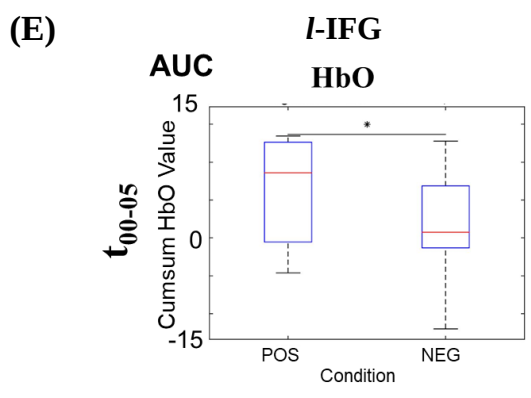


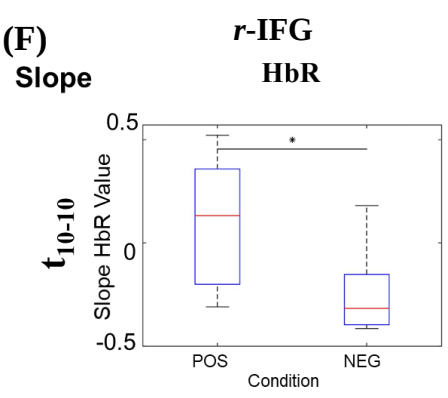


		HbO	HbT	HbR
Average	t_{05-15}	0.06	0.14	0.45
	t_{05-10}	0.12	0.21	0.65
	t_{10-20}	0.18	0.85	0.21
	t_{10-15}	**0.01**	**0.04**	0.33
	t_{00-20}	0.14	0.54	0.41
AUC	t_{05-15}	0.10	0.20	0.57
	t_{05-10}	0.21	0.36	0.67
	t_{10-20}	**0.03**	0.15	0.29
	t_{10-15}	**0.01**	**0.03**	0.40
	t_{00-20}	0.14	0.45	0.57
Slope	t_{00-15}	0.88	0.56	0.47
	t_{00-10}	0.93	0.63	0.29
	t_{00-05}	0.93	0.60	0.17

		HbO	HbT	HbR
Average	t_{05-15}	0.26	0.25	0.84
	t_{05-10}	0.27	0.24	0.73
	t_{10-20}	0.26	0.30	0.93
	t_{10-15}	0.30	0.31	0.94
	t_{00-20}	0.19	0.20	0.97
AUC	t_{05-15}	0.25	0.23	0.78
	t_{05-10}	0.21	0.17	0.70
	t_{10-20}	0.29	0.31	0.98
	t_{10-15}	0.34	0.35	0.89
	t_{00-20}	0.17	0.16	0.83
Slope	t_{00-15}	**0.02**	**0.03**	0.62
	t_{00-10}	**0.02**	**0.02**	0.90
	t_{00-05}	**0.03**	**0.02**	0.84

		HbO	HbT	HbR
Average	t_{05-15}	**0.01**	**0.02**	0.40
	t_{05-10}	**0.01**	**0.02**	0.57
	t_{10-20}	0.07	0.08	0.29
	t_{10-15}	**0.04**	**0.04**	0.31
	t_{00-20}	**0.03**	**0.05**	0.32
AUC	t_{05-15}	**0.01**	**0.03**	0.50
	t_{05-10}	**0.02**	**0.04**	0.63
	t_{10-20}	**0.04**	**0.05**	0.29
	t_{10-15}	**0.03**	**0.03**	0.33
	t_{00-20}	**0.02**	**0.05**	0.42
Slope	t_{00-15}	0.60	0.10	0.91
	t_{00-10}	0.54	0.09	0.79
	t_{00-05}	0.25	0.17	0.99



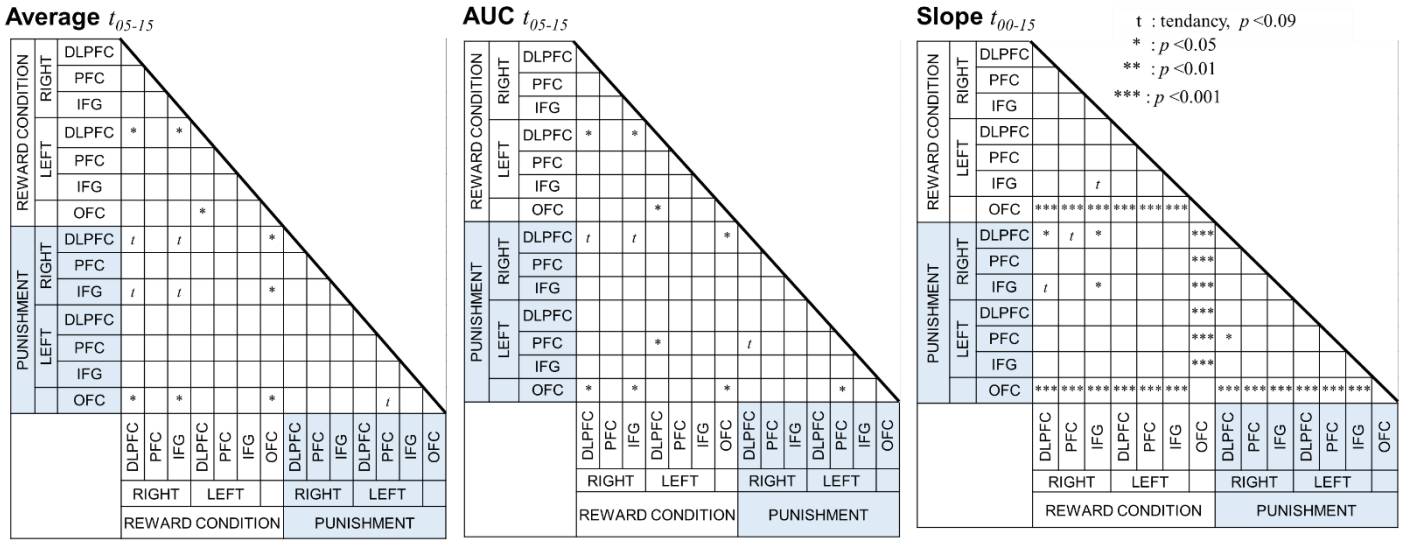


		HbO	HbT	HbR
Average	t_{00-05}	**0.06**	0.56	0.59
	t_{00-10}	0.25	0.96	0.94
AUC	t_{00-05}	**0.05**	0.45	0.56
	t_{00-10}	0.10	0.72	0.75
Slope	t_{00-05}	0.39	0.90	0.48
	t_{00-10}	0.38	0.51	0.55

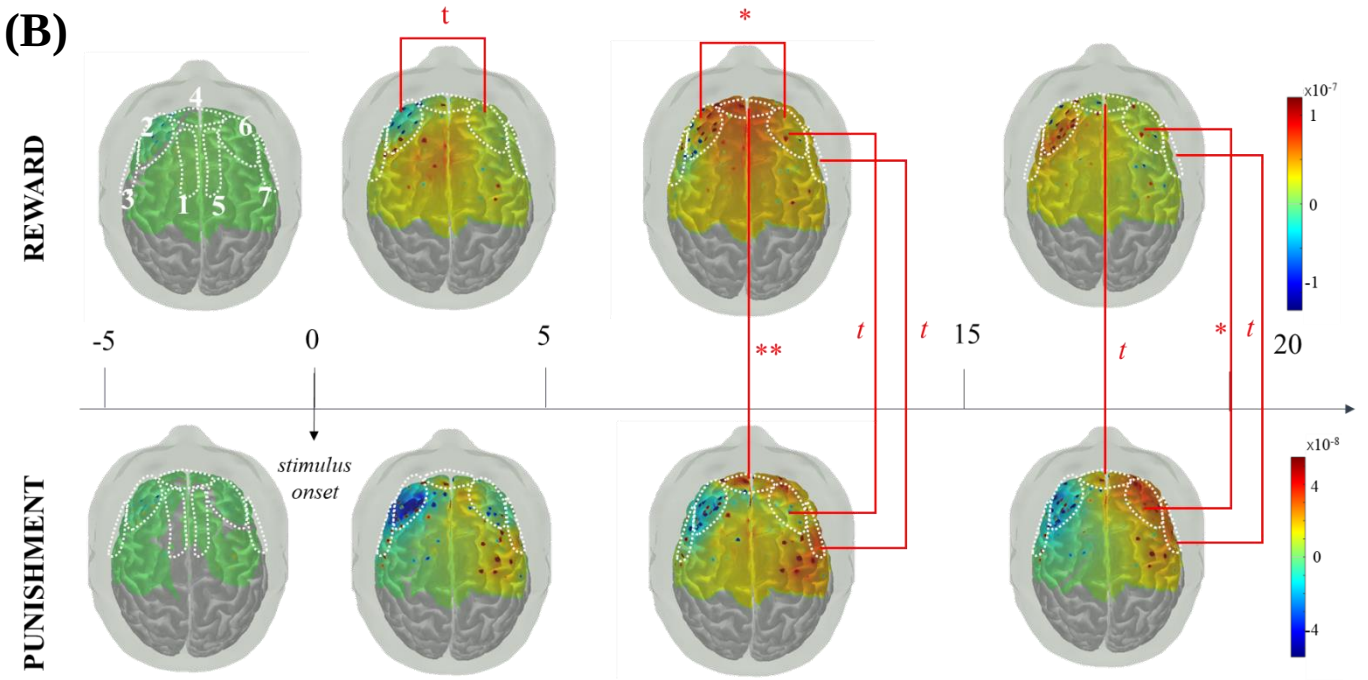
		HbO	HbT	HbR
Average	t_{00-05}	0.17	0.24	**0.08**
	t_{00-10}	0.21	0.27	**0.09**
AUC	t_{00-05}	0.21	0.31	**0.09**
	t_{00-10}	0.18	0.24	**0.08**
Slope	t_{00-05}	0.12	0.09	**0.07**
	t_{00-10}	0.16	0.23	**0.02**

Figure 5

(A)



(B)



(C)

		Reward	Punishment	Reward vs Punishment	Post Hoc Reward vs Punishment
Amplitude	**l-DLPFC**	Significant (HbO)		Significant (HbO & HbT)	Significant vs r-DLPFC under *Reward* (HbO)
	l-IFG	Significant (HbO)		Tendency (HbO)	
	r-DLPFC		Tendency (HbO)		Tendency (HbO) + Significant vs l-DLPFC under *Reward* (HbO)
	r-IFG		Tendency (HbO)	Tendency (HbR)	Tendency (HbO)
	OFC		Significant (HbO)	Significant (HbO & HbT)	Significant (HbO)
AUC	**l-DLPFC**	Significant (HbO)		Significant (HbO & HbT)	Significant vs r-DLPFC under *Reward* (HbO)
	l-IFG	Significant (HbO)		Tendency (HbO)	
	r-DLPFC		Tendency (HbO)		Tendency (HbO) + Significant vs l-DLPFC under *Reward* (HbO)
	r-IFG		Tendency (HbO)	Tendency (HbR)	
	OFC		Significant (HbO)	Significant (HbO & HbT)	Significant (HbO)
Slope	**l-DLPFC**				
	l-IFG				
	r-DLPFC			Significant (HbO & HbT)	Significant (HbO)
	r-IFG			Tendency (HbR)	Significant (HbO)
	OFC				Significant (HbO)