Cortical Hemodynamic Mechanisms of Reversal learning using High-Resolution Functional Near-Infrared Spectroscopy: A Pilot Study

charlotte piau

INSERM U1105: GRAMFC

Mahdi Mahmoudzadeh (✉ mahdi.mahmoudzadeh@gmail.com)

Centre Hospitalier Universitaire D' Amiens <https://orcid.org/0000-0003-3145-2133>

Astrid Kibleur

University Grenoble Alpes: Universite Grenoble Alpes

Mircea Polosan

University Grenoble Alpes: Universite Grenoble Alpes

Olivier David

University Grenoble Alpes: Universite Grenoble Alpes

Farbice Wallois

INSERM U1105: GRAMFC

Research

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

Posted Date: April 6th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-379282/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: 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. Here, in 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: These results 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.

Highlights

- We detected hemispheric asymmetry of functional frontal brain activity between punishment and reward processing following reversal of a learned rule.
- The left dorsolateral prefrontal cortex and inferior frontal gyrus were significantly activated under the reward condition.
- The orbito-frontal cortex was significantly activated under the punishment condition and we observed a tendency towards activation for the right dorsolateral prefrontal and inferior frontal areas.

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) (Ghahremani et al., 2010). Paradigms such as reversal-learning tasks (RLTs) can be used to measure behavioral flexibility (Chamberlain et al., 2008; Ghahremani et al., 2010; Izquierdo et al., 2017; Xue et al., 2013). 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 (Ghahremani et al., 2010). Impairments in reversal-learning processes are associated with a wide range of abnormal behavioral abnormalities and psychiatric conditions characterized by impulsiveness and disinhibition, such as obsessional compulsive disorder (OCD) (Chamberlain et al., 2008).

Lesional studies have been performed to characterize the underlying neural mechanisms that are solicited in the RLT paradigm (Dias et al., 1996; Izquierdo et al., 2004; *Fellows et Farah., 2003*; Hornak et al., 2004) and have consistently reported involvement of the ventrolateral prefrontal cortex (VLPFC) and lateral orbitofrontal cortex (OFC). In addition, functional magnetic resonance imaging (fMRI) studies have highlighted the involvement of other brain regions, including the lateral OFC (*Hampshire et al., 20012*; Chamberlain et al., 2008; O'Doherty et al., 2001), the inferior frontal gyrus (IFG) (Cools et al., 2002; Nagahama et al., 2001), the dorsomedial prefrontal cortex (DMPFC) (Budhani et al., 2007), the dorsolateral prefrontal cortex (DLPFC) (Remijnse et al., 2005; Ghahremani et al., 2010), the posterior parietal cortex (Hampshire and Owen, 2006), the anterior cingulate cortex (ACC) (Chudasama et al., 2013; Kawai et al., 2015; Izquierdo et al., 2017), and the striatum (Cools et al., 2002; *Hampton et O'Doherty, 2007*). Recently, several studies examined how reward and punishment can modulate reversal learning (Xue et al., 2013). Reward and punishment represent two major motivations to learn in changing environments. Several RLT studies have shown the involvement of the left frontal hemisphere with reward feedback (O'Doherty et al., 2001; O'Doherty et al., 2003; Remijnse et al., 2005). Others have reported frontal lateralization according to the valence of the feedback. Positive outcomes are lateralized to the left frontal regions (Sobotka et al., 1992; *Bechara et Damasio 2005*; Xue et al., 2009), whereas unpleasant feedback involves the right frontal hemisphere.

Functional near infra-red spectroscopy (fNIRS) is used non-invasively in a natural environment for human infants and adults to analyze cortical activation (Wilcox and Biondi, 2015; León-Carrión and León-Domínguez, 2012, Mahmoudzadeh et al., 2013) or in pathological situations (*Roche Labarbe et al., 2008*). Brain activation is associated with neurovascular coupling that induces hemodynamic changes, modifying the optical properties of brain tissue, which can be assessed by fNIRS (León-Carrión and León-Domínguez, 2012, *Jobsis et al., 1977*). Here, we used the powerful temporal resolution of fNIRS (Cui et al., 2011), 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 observe 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 confirm the involvement of the DLPFC, IFG, and OFC in a RLT, as observed by fMRI (Nagahama et al., 2001; Cools et al., 2002; Remijnse et al., 2005; Ghahremani et al., 2010; Xue et al., 2013; Izquierdo et al., 2017). Thus, we aimed to characterize the involvement of the DLPFC and the IFG during the reversal process and differentiate their activity under reward and punishment learning conditions using an innovative noninvasive portable solution using high density fNIRS. To date, no one has studied the RLT paradigm using this approach, and replicating some of the fMRI findings using fNIRS is of particular

translational value for future studies in subjects who cannot be recorded in fMRI, e.g. subjects with brain electrodes.

Materials And Methods

Participants

Eight healthy right-handed participants (6 women and 2 men) aged between 22 and 56 years were enrolled in the study. All subjects had normal or corrected to normal vision and 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.

Task

During the 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. Each process block had a duration of 360 s and was composed of 80 randomized stimuli (40 positive and 40 negative). The organization of the task is shown in Fig. 1. Every trial began with presentation of a pair of animal pictures from the IAPS (Lang et al., 2008), of neutral valence and intensity, presented on a grey background for a maximum of 300 ms. Each trial was followed by a fixation-cross interstimulus lasting for 1000–1500 ms (randomly determined). The positions of the pictures of animals were randomized (one on top, the other at the bottom of the screen). During the 300 ms, participants had to decide behind which animal money was hidden. Once the response was made by the participant, feedback was delivered (Fig. 1). 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 Fig. 1 about here]

Behavioral responses

Behavioral responses were labelled according to Reminsje et al. (*Reminsje et al., 2005*) (Fig. 1). 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 a 32-detector and 16-emitter design (Fig. 2A). Optodes covered the prefrontal, frontal, and parietal lobes, with an emitter-detector distance of 2.5 cm. 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) (Fig. 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 (*Aasted et al., 2005*) (Fig. 2A and 2B). The fNIRS sensitivity map (Fig. 2C) was modelled using AtlasViewer and freely available Monte-Carlo photon transport software (tMCimg routine, number of simulated photons = 10^6). The more optimal sensitivity configuration of the fNIRS source and detector positions is presented in Fig. 2A.

[Insert Fig. 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 *accuracy* (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. The number of successive errors was also evaluated. These errors, called *perseverative reversal errors*, indicate that the former discrimination strategy, which is now obsolete, was still used, indicating that the participant had not yet updated his/her strategy.

Several studies have evaluated the effect of reward and punishment feedback on reversal learning by focusing on contrasting reversal learning with the acquisition phase (*Reminsje et al., 2005*; Budhani et al., 2007; Ghahremani et al., 2010; *Xue et al., 2017*). Hemodynamic analysis was performed according to Reminsje (2002). 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) (Fig. 1). Such a contrast was performed here for the reward and punishment conditions. Contrasts were performed at the single-subject level.

Homer2 Matlab toolbox was used to analyze the fNIRS signal (Huppert et al., 2009). A band-pass filter between 0.03 and 0.1 Hz was applied to eliminate physiological noise (very low-frequency oscillations, respiration, and heartbeat). A window from – 5 to 30 s around the onset of the stimulation ($t = 0$ s) was used to analyze the hemodynamic response to RLT (Buxton, 2001). The fixation-cross interstimulus interval throughout the whole task, served as the implicit baseline to allow the estimation of the hemodynamic response function by simple averaging (Costantini et al., 2013).

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 homogenize the data according to the individual characteristics. A baseline correction [-5, 0] s was finally applied to the normalized data.

Region of interest (ROI)

We were able to associate each channel with a cortical ROI using the AtlasViewer toolbox (*Aasted et al., 2005*) 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) (Fig. 2D). Channels with an inter-optode distance < 5 cm were then selected (León-Carrión

and León-Domínguez, 2012). 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 (*Smielewski 1995*).

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 subject 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 t_0 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 (Mandrick et al., 2013). 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). (*Aasted et al, 2005*).

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 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-test, 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] sec) 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 all comparison, we only report the significant results ($p < 0.05$) and 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 Fig. 3. We used t-test to compare the following accuracy and reaction time parameters. During the acquisition phase (Fig. 3A – *left panel*), between the first and following presentations (2–8), the accuracy increased non-significantly (Fig. 3A – *left panel*) for the reward (Accuracy presentation 1 vs 2–8 + 17%, $p > 0.1$) and punishment (Accuracy presentation 1 vs 2–8 +29%, $p = 0.05$) conditions. We observed no significant differences for the averaged accuracy between the punishment and reward conditions ($p > 0.1$).

The reaction time (RT) (Fig. 3B – *left panel*) decreased significantly between the first and second presentation under both the punishment (RT presentation 1 vs 2 : -0,4 sec, $p < 0.05$) and reward RT presentation 1 vs 2 : -0,5 sec, $p < 0.01$) conditions. We observed no significant differences between the two conditions ($p > 0.1$) for the averaged RT.

During the reversal phase (Fig. 3A – *right panel*), between the first and following presentations (2–8), the accuracy (Fig. 3A – *right panel*) increased to 69% for the punishment (presentation 1 vs 2–8, $p < 0.001$) and 57% for the reward conditions (presentation 1 vs 2–8 $p < 0.001$), 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 presentation 1 vs 2: +46% for the reward ($p = 0.002$) and 66% for the punishment ($p < 0.001$) conditions, respectively]. The differences for the averaged accuracy between the reward and punishment conditions ($p < 0.01$) were significant for all presentations. For RT (Fig. 3B – *right panel*), we observed a non-significant increase between presentation 1 and presentation 2 for the punishment (-0,06 sec) and reward (-0,08 sec) conditions ($p > 0.05$). Comparison of the averaged reaction time for all presentations showed the difference between the two conditions to be significant ($p < 0.001$).

Finally, the increase in the number of perseverative errors (Fig. 3C) between the reward and punishment conditions was not significant ($p = 0.07$), 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 more negatively affected the reward condition than the punishment condition.

Cortical hemodynamic response

We analyzed the detection of the unexpected outcome by computing the hemodynamic event contrast as described by Reminsje, (2002): (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 (Ferrari and Quaresima, 2012; Perrey et al., 2010).

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) using t-test. We observed significant activation for the left hemisphere for both conditions (Fig. 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 ($p < 0.036$), with a latency to the peak of approximately 7 s (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 ($p < 0.045$), with a similar pattern and latency to the peak (approximately 3 sec, 1.3 AU, $c = 0.21$). There was a second peak at approximately 12 s (0.8 AU). 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, $p < 0.075$; r-IFG: 1.09AU, $c = 0.20$, with $t = [2.5, 4]$ s, $p < 0.075$) 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 sec (1.42 AU, with $t = [12, 15.5]$ s, $p < 0.075$). Finally, the OFC showed a significant increase in HbO, with a latency to the peak of approximately 13 s (1.45 AU, $c = 0.11$).

T-test comparison of the AUC for the 5-s subperiods to that of the baseline showed involvement of the l-DLPFC ($p < 0.01$ t_{05-15}) and l-IFG ($p < 0.05$ t_{00-15}) under the reward condition. Under the punishment condition, we also observed the involvement of the OFC ($p < 0.05$ t_{00-15}) and a tendency towards involvement of the r-DLPFC ($p < 0.08$ for t_{03-08}) and r-IFG ($p < 0.08$ for t_{00-05} and t_{10-15}).

We then evaluated these differences in activation between the two conditions (reversal vs punishment) for the same ROI using a student t-test, taking into account the three parameters (slope, AUC, average) across the subperiods.

Statistical analysis (Fig. 4B) of the changes in HbO in the l-DLPFC showed a significant difference between conditions for the AUC ($p < 0.02$) and average ($p < 0.02$) for the period t_{10-15} . There were also significant differences for HbT (AUC, $p < 0.03$; average, $p < 0.04$). We observed no significant differences in any parameters for HbR. For the r-DLPFC, we observed significant differences between conditions for HbO (slope, $p < 0.03$) (Fig. 4C) and HbT (Slope $p < 0.03$), but none for HbR or any other parameters.

The r-IFG showed significant differences between conditions for HbR for t_{00-10} (slope, $p < 0.021$; AUC, $p < 0.08$; average, $p < 0.09$), but none for HbO or HbT (Fig. 4F). The l-IFG showed a tendency for HbO for t_{00-05} (average, $p < 0.06$; AUC, $p < 0.05$) (Fig. 4E). We also observed significant differences for various subperiods (t_{05-15} , t_{05-10} , t_{10-20} , t_{10-15} and t_{00-20}) for HbO and HbT in the OFC (Fig. 4D).

Brain regions differentially affected by reversal by the effect of ROI $\times$ 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-DLPFC, R-IFG, L-PFC, L-DLPFC, L-IFG)* for the three previously identified significant subperiods (*Table S1*). We thus observed a tendency towards a significant effect of the *ROI × condition* on HbO (Average t_{05-15} $F_{(6,98)} = 1.93$, $p < 0.07$; AUC t_{05-15} $F_{(6,98)} = 2.00$, $p < 0.07$) for t_{05-15} . We also observed a tendency towards a significant effect of *condition* (*Table S1.A*) for the HbR (slope t_{00-15} $F_{(1,98)} = 3.22$, $p < 0.08$) and a significant effect for HbO (Average t_{10-20} $F_{(1,98)} = 5.26$, $p = 0.024$; AUC t_{10-20} $F_{(1,98)} = 4.53$, $p = 0.035$) and HbT (AUC t_{10-20} $F_{(1,98)} = 5.88$, $p = 0.017$; Average t_{10-20} $F_{(1,98)} = 6.25$, $p < 0.014$). An effect of the *ROI* (*Table S1.B*) was observed for the HbO slope parameters (slope t_{00-15} $F_{(6,98)} = 19.32$, $p < 0.001$).

We then performed a Fisher post-hoc analysis to define the ROIs and conditions that showed a significant effect. This analysis was first performed for the t_{05-15} period and if no significant results were observed, it was performed for the t_{05-10} and t_{10-15} periods. For clarity, we present the results in three cross-tables, one for each parameter (slope, AUC, and average) (Fig. 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 (Fig. 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, $p < 0.05$; AUC, $p < 0.05$) and the OFC between the two conditions (AUC, $p < 0.05$; average, $p < 0.07$). There was also a significant difference for the slope parameters between the two conditions for the r-DLPFC and r-IFG ($p < 0.05$).

The statistical analysis is added to the HbO hemodynamic mapping in Fig. 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 ($p < 0.08$).

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 (Fig. 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 ($p < 0.08$) towards a typical hemodynamic response was observed over the r-DLPFC and r-IFG. These observations were confirmed by statistical and post-hoc analysis.

Discussion

Previous studies on reversal learning used fMRI (O'Doherty et al., 2001; Cools et al., 2002; Remijne et al., 2005; Ghahremani et al., 2010; Xue et al., 2013) or EEG (Sobotka et al., 1992) tools. In this study, we investigate the cortical hemodynamic response due to a reversal-learning task with fNIRS tool. fNIRS has a better temporal resolution, relative to fMRI (Cui et al., 2011), and better spatial resolution, relative to EEG (Parasuraman and Caggiano, 2005). Thus, fNIRS emerges as an alternative tool for subjects who cannot undergo fMRI recordings.

Here, we reveal distinct neural substrates for reversal learning driven by reward and punishment using comparable magnitude feedback, as proposed by Xue et al., 2013. 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 significance was observed for the r-DLPFC and r-IFG.

Behavioral data from our study show that all subjects understood the rules, as they succeeded in accumulating a positive amount 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 ($p < 0.02$) for the l-DLPFC for t_{10-15} 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_{00-10} in terms of the HbR parameter under the punishment condition.

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

The DLPFC has been examined in previous neuroimaging studies on RLT, either bilaterally (Chamberlain et al., 2008; Cools et al., 2002; Waegeman et al., 2014) or only the left (*Fellow and Farah, 2003*) or right hemisphere (O'Doherty et al., 2001; Xue et al., 2013). Functional evidence has also shown the involvement of the OFC in reversal reward processing (Delgado et al., 2000; Cools et al., 2002; Izquierdo et al., 2017), more precisely the l-OFC (*Sobotka et al., 1998; Bechara et al., 2005; Xue et al., 2009*), whereas the r-OFC has been linked to punishment processing (Sobotka et al., 1992; O'Doherty et al., 2001). Concerning reinforcement learning, Xue et al. revealed distinct mechanisms underlying learning from positive and negative feedback (i.e. reward and punishment, respectively). They observed rightward asymmetry (right lateral OFC and DLPFC) in punishment processing under punishment conditions, but nothing was observed for the reward condition. In this study, Xue et al., used different types of feedback (shock and money) associated with an incomparable level of magnitude between the two conditions. As suggested by Xue et al., we used comparable feedback (gain or loss of money) with an equal level of magnitude (+ 100\$ vs. -100\$) to investigate local asymmetries in the treatment of information in RLT

specific to each condition (reward and punishment). The prefrontal cortex (PFC) regulates and monitors a number of "executive" cognitive functions (Weinberger, 1993), whereas the DLPFC is involved in various cognitive tasks: working memory (Belger et al., 1998), reasoning (*Prabhakaran et al., 1997*), changes in attention (Dias et al., 1996), and control (*Hampshire et Owen, 2006*). In accordance with these previous results, neurovascular coupling was observed over the l-DLPFC under the reward condition during the RLT. However, as reported in other studies (O'Doherty et al., 2001; Cools et al., 2002; Xue et al., 2013), our results also suggest the possible involvement of the r-DLPFC under punishment conditions. These two observations support the involvement of the l-DLPFC in the detection of switching (Remijne et al., 2005; Cools et al., 2002; Xue et al., 2013) and the updating of response-outcome relationships and flexible behavior (Ghahremani et al., 2010; *O'Doherty et al., 2003*). Moreover, its involvement is lateralized according to the condition: l-DLPFC for the reward condition and r-DLPFC for the punishment condition.

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

The IFG is involved in inhibitory control (Kawashima et al., 1996; Konishi et al., 1998; Konishi et al., 1999; Garavan et al., 1999; de Zubicaray et al., 2000; Swick et al., 2008; *Rygula et al., 2010*; Aron et al., 2014). The IFG has also been shown to be involved in the RLT (Hampshire and Owen, 2006; Budhani et al., 2007), mostly in the right hemisphere (Cools et al., 2002; Ghahremani et al., 2010; Waegeman et al., 2014). 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. This result confirms the role of the IFG in inhibiting a well-learned association (Ghahremani et al., 2010; Aron et al., 2004, Rygula et al., 2010) 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., (2001) 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 (*O'Doherty et Dolan, 2006*; Wallis, 2007) and motivation-related processes (Rothkirch et al., 2012; Spielberg et al., 2012). More precisely, the role of the OFC in reversal is to store the feedback association during reversal learning (*Cai et Padoa-Schioppa, 2014*; Keiflin et al., 2013; *Moorman et Aston-Jones, 2014*). 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 (*Davidson et Irwin, 1999*; *Deglin et Kinsbourne, 1996*; Overskeid, 2000; *Bechara et Damasio 2005*) and positive affects (Baxter et al., 1989; *Davidson et Henriques, 2000*; *Harmon-Jones et Allen, 1997*; *Herrington et al., 2007*). 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 (Sobotka et al., 1992; Xue et al., 2009). 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 (Szatkowska et al., 2008), which could then be used to select appropriate goals. This relationship is bidirectional (Spielberg et al., 2012). Indeed, the involvement of the I-DLPFC (I-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 (*Spielber et al., 2012; Hare et al., 2009*).

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 (*Dosenbach et al., 2007*). One involved the DLPFC, which is associated with top-down attentional control (*Dosenbach et al., 2007*), maintaining goals, and updating information (*Wager et Smith, 2003*), whereas the other involved the anterior cingulate cortex (ACC), which is involved in detecting conflicting responses and monitoring performance (Nelson et al., 2003; Seeley et al., 2007; Cole et al., 2013; Banich, 2019). In our study, involvement of the I-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 (Duann et al., 2009; Neubert et al., 2010; Swann et al., 2012, Spechler et al., 2016) but studies have also suggested involvement of the subthalamic nucleus (STN), which is linked to the motor inhibition loop (*Aron et Poldrack, 2006; Schmidt et al., 2013*). 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 (*Aron, 2011; Bari et Robbins, 2013, Verbruggen et Logan, 2008; Cunillera et al., 2014; Tops et al., 2014; Cai et al., 2016*). In addition, the involvement of the IFG in our study may be indicative of its participation in this inhibitory loop (*Jonides et Nee, 2006; Swick et al., 2008*). Consistent with this hypothesis, the results concerning the slope coefficient suggest 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, the DLPFC recruits the IFG via an excitatory connection to initiate the longer indirect inhibitory pathway. Second, the IFG may inhibit DLPFC activity during reappraisal once the strategy process is updated (*Banich et al., 2019; Morawetz et al., 2016*). This second activation of the IFG may explain the second peak observed for the r-IFG under the punishment condition and the l-IFG under the reward condition.

In summary, we used fNIRS to confirm the differential impact of reversal: with involvement of the left hemisphere under the reward-guided condition, characterized by early I-DLPFC activation followed by

involvement of the I-FG and OFC, and right hemisphere involvement (r-DLPFC and r-IFGC) under the punishment condition. This suggests feedback-specific processing during RLTs.

One of the limitations of this study had to do with the small number of participants. Indeed, the present study was a preliminary analysis before taking into account participants who had undergone deep brain stimulation.

Conclusion

This study extends our understanding of the neural mechanisms involved in reversal learning and improves our comprehension of risky behaviors in vulnerable populations characterized by impaired flexibility. This approach demonstrates the possibility of using high-density optical imaging tools to study cognitive tasks and can provide insights into neurophysiological mechanisms and facilitate the translation of functional optical imaging into clinical applications, such as in OCD or Parkinson disease.

Declarations

Ethics approval and consent to participate

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.

Consent for publication

The corresponding author, on behalf all authors, allows the publication of this Manuscript.

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

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.

Authors' contributions

All authors contributed to conception and design of the study. MM designed the fNIRS setup. CP and MM participates in data acquisition. CP and MM analyzed fNIRS data. AK analyzed behavior data. CP and MM performed the statistical analysis. CP, MM, AK, and FW wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Acknowledgements

Not applicable.

References

- Aasted CM, Yücel MA, Cooper RJ, Dubb J, Tsuzuki D, Becerra L, Petkov MP, Borsook D, Dan I, Boas DA. Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. *Neurophotonics*. 2015;2(2):020801. <https://doi.org/10.1117/1.NPh.2.2.020801>.
- Aron AR. From reactive to proactive and selective control: Developing a richer model for stopping inappropriate responses. *Biol Psychiat*. 2011;69(12):e55–68. <https://doi.org/10.1016/j.biopsych.2010.07.024>.
- Aron AR, Poldrack RA. Cortical and Subcortical Contributions to Stop Signal Response Inhibition: Role of the Subthalamic Nucleus. *The Journal of Neuroscience*. 2006;26(9):24242433. <https://doi.org/10.1523/JNEUROSCI.4682-05.2006>.
- Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*. 2004;8(4):170177. <https://doi.org/10.1016/j.tics.2004.02.010>.
- Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*. 2014;18(4):177185. <https://doi.org/10.1016/j.tics.2013.12.003>.
- Banich MT. (2019). The Stroop Effect Occurs at Multiple Points Along a Cascade of Control: Evidence From Cognitive Neuroscience Approaches. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02164>.
- Bari A, Robbins TW. Inhibition and impulsivity: Behavioral and neural basis of response control. *Prog Neurobiol*. 2013;108:4479. <https://doi.org/10.1016/j.pneurobio.2013.06.005>.
- Baxter LR, Schwartz JM, Phelps ME, Mazziotta JC, Guze BH, Selin CE, Gerner RH, Sumida RM. Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Arch Gen Psychiatry*. 1989;46(3):243250. <https://doi.org/10.1001/archpsyc.1989.01810030049007>.
- Bechara A, Damasio AR. The somatic marker hypothesis: A neural theory of economic decision. *Games Economic Behavior*. 2005;52(2):336372. <https://doi.org/10.1016/j.geb.2004.06.010>.

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(1):1432.

Bode S, Haynes J-D. Decoding sequential stages of task preparation in the human brain. *NeuroImage*. 2009;45(2):606613. <https://doi.org/10.1016/j.neuroimage.2008.11.031>.

Budhani S, Marsh AA, Pine DS, Blair RJR. Neural correlates of response reversal: Considering acquisition. *NeuroImage*. 2007;34(4):17541765. <https://doi.org/10.1016/j.neuroimage.2006.08.060>.

Buxton RB. The Elusive Initial Dip. *NeuroImage*. 2001;13(6):953958. <https://doi.org/10.1006/nimg.2001.0814>.

Cai X, Padoa-Schioppa C. Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron*. 2014;81(5):11401151. <https://doi.org/10.1016/j.neuron.2014.01.008>.

Cai Y, Li S, Liu J, Li D, Feng Z, Wang Q, Chen C, Xue G. 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(1):177186. https://doi.org/10.1162/jocn_a_00888.

Chamberlain SR, Menzies L, Hampshire A, Suckling J, Fineberg NA, del Campo N, Aitken M, Craig K, Owen AM, Bullmore ET, Robbins TW, Sahakian BJ. Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. *Science*. 2008;321(5887):421422. <https://doi.org/10.1126/science.1154433>.

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. *Cereb Cortex*. 2013;23(12):28842898. <https://doi.org/10.1093/cercor/bhs266>.

Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature neuroscience*. 2013;16(9):13481355. <https://doi.org/10.1038/nn.3470>.

Cools R, Clark L, Owen AM, Robbins TW. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 22(11), 45634567. <https://doi.org/20026435>.

Costantini M, Vacri AD, Chiarelli AM, Ferri F, Romani GL, Merla A. Studying social cognition using near-infrared spectroscopy: The case of social Simon effect. *J Biomed Opt*. 2013;18(2):025005. <https://doi.org/10.1117/1.JBO.18.2.025005>.

Cui X, Bray S, Bryant DM, Glover GH, Reiss AL. A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *NeuroImage*. 2011;54(4):28082821.

<https://doi.org/10.1016/j.neuroimage.2010.10.069>.

Cunillera T, Fuentemilla L, Brignani D, Cucurell D, Miniussi C. (2014). A Simultaneous Modulation of Reactive and Proactive Inhibition Processes by Anodal tDCS on the Right Inferior Frontal Cortex. PLoS ONE, 9(11). <https://doi.org/10.1371/journal.pone.0113537>.

Davidson RJ, Henriques J. (2000). Regional brain function in sadness and depression. In *The neuropsychology of emotion* (p. 269-297). Oxford University Press.

Davidson RJ, Irwin W. The functional neuroanatomy of emotion and affective style. Trends in Cognitive Sciences. 1999;3(1):1121. [https://doi.org/10.1016/S1364-6613\(98\)01265-0](https://doi.org/10.1016/S1364-6613(98)01265-0).

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(9):1280-1291. [https://doi.org/10.1016/S0028-3932\(00\)00033-6](https://doi.org/10.1016/S0028-3932(00)00033-6).

Deglin VL, Kinsbourne M. Divergent Thinking Styles of the Hemispheres: How Syllogisms Are Solved during Transitory Hemisphere Suppression. Brain Cogn. 1996;31(3):285-307. <https://doi.org/10.1006/brcg.1996.0048>.

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(6):3072-3077. <https://doi.org/10.1152/jn.2000.84.6.3072>.

Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. Nature. 1996;380(6569):6972. <https://doi.org/10.1038/380069a0>.

Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA. 2007;104(26):11073-11078. <https://doi.org/10.1073/pnas.0704320104>.

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. The Journal of Neuroscience. 2009;29(32):10171-10179. <https://doi.org/10.1523/JNEUROSCI.1300-09.2009>.

Fellows LK, Farah MJ. Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal learning paradigm. Brain: A Journal of Neurology. 2003;126(Pt 8):1830-1837. <https://doi.org/10.1093/brain/awg180>.

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(2):921-935. <https://doi.org/10.1016/j.neuroimage.2012.03.049>.

- 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(14):8301-8306. <https://doi.org/10.1073/pnas.96.14.8301>.
- Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA. (2010). Neural Components Underlying Behavioral Flexibility in Human Reversal Learning. *Cerebral Cortex (New York, NY)*, 20(8), 1843-1852. <https://doi.org/10.1093/cercor/bhp247>.
- 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(4):4102-4112. <https://doi.org/10.1016/j.neuroimage.2011.10.072>.
- Hampshire A, Owen AM. (2006). Fractionating attentional control using event-related fMRI. *Cerebral Cortex (New York, N.Y.: 1991)*, 16(12), 1679-1689. <https://doi.org/10.1093/cercor/bhj116>.
- Hampton AN, O'Doherty JP. Decoding the neural substrates of reward-related decision making with functional MRI. *Proc Natl Acad Sci USA*. 2007;104(4):1377-1382. <https://doi.org/10.1073/pnas.0606297104>.
- Hare TA, Camerer CF, Rangel A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*. 2009;324(5927):646-648. <https://doi.org/10.1126/science.1168450>.
- 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(1):159-163. <https://doi.org/10.1037//0021-843x.106.1.159>.
- Herrington JD, Mohanty A, Koven NS, Fisher JE, Stewart JL, Banich MT, Webb AG, Miller GA, Heller W. Emotion-modulated performance and activity in left dorsolateral prefrontal cortex. *Emotion*. 2005;5(2):200-207. <https://doi.org/10.1037/1528-3542.5.2.200>.
- Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J Cogn Neurosci*. 2004;16(3):463-478. <https://doi.org/10.1162/089892904322926791>.
- 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(10):D280-98.
- Izquierdo A, Brigman JL, Radke AK, Rudebeck PH, Holmes A. The neural basis of reversal learning: An updated perspective. *Neuroscience*. 2017;345:1226. <https://doi.org/10.1016/j.neuroscience.2016.03.021>.
- Izquierdo A, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2004;24(34):7540-7548. <https://doi.org/10.1523/JNEUROSCI.1921-04.2004>.

Jöbsis FF. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*. 1977;198(4323):12641267.

Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. *Neuroscience*. 2006;139(1):181193. <https://doi.org/10.1016/j.neuroscience.2005.06.042>.

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(4):792804. <https://doi.org/10.1016/j.neuron.2015.09.030>.

Kawashima R, Satoh K, Itoh H, Ono S, Furumoto S, Gotoh R, Koyama M, Yoshioka S, Takahashi T, Takahashi K, Yanagisawa T, Fukuda H. Functional anatomy of GO/NO-GO discrimination and response selection—A PET study in man. *Brain Res*. 1996;728(1):7989.

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(40):1598915998. <https://doi.org/10.1523/JNEUROSCI.0026-13.2013>.

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: A Journal of Neurology*. 1999;122 (Pt(5):981991. <https://doi.org/10.1093/brain/122.5.981>.

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. *The European Journal of Neuroscience*. 1998;10(3):12091213. <https://doi.org/10.1046/j.1460-9568.1998.00167.x>.

Lang PJ, Bradley MM, Cuthbert BN. (2008). **International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.**

León-Carrión J, León-Domínguez U. Functional Near-Infrared Spectroscopy (fNIRS): Principles and Neuroscientific Applications. *Neuroimaging - Methods*. 2012. <https://doi.org/10.5772/23146>.

Mahmoudzadeh M, Dehaene-Lambertz G, Fournier M, Kongolo G, Goudjil S, Dubois J, Grebe R, Wallois F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, 110(12), 48464851. <https://doi.org/10.1073/pnas.1212220110>.

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. *Int J Ind Ergon*. 2013;43(4):335341. <https://doi.org/10.1016/j.ergon.2013.05.003>.

Moorman DE, Aston-Jones G. Orbitofrontal Cortical Neurons Encode Expectation-Driven Initiation of Reward-Seeking. *J Neurosci*. 2014;34(31):1023410246. <https://doi.org/10.1523/JNEUROSCI.3216-13.2014>.

- Morawetz C, Bode S, Baudewig J, Kirilina E, Heekeren HR. (2016). Changes in Effective Connectivity Between Dorsal and Ventral Prefrontal Regions Moderate Emotion Regulation. *Cerebral Cortex (New York, N.Y.: 1991)*, 26(5), 1923-1937. <https://doi.org/10.1093/cercor/bhv005>.
- Nagahama Y, Okada T, Katsumi Y, Hayashi T, Yamauchi H, Oyanagi C, Konishi J, Fukuyama H, Shibasaki H. (2001). Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 11(1), 85-92. <https://doi.org/10.1093/cercor/11.1.85>.
- Nelson JK, Reuter-Lorenz PA, Sylvester C-YC, Jonides J, Smith EE. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proceedings of the National Academy of Sciences*, 100(19), 11171-11175. <https://doi.org/10.1073/pnas.1334125100>.
- 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 USA*. 2010;107(30):13240-13245. <https://doi.org/10.1073/pnas.1000674107>.
- 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(1):95-102. <https://doi.org/10.1038/82959>.
- O'Doherty J, Critchley H, Deichmann R, Dolan RJ. Dissociating Valence of Outcome from Behavioral Control in Human Orbital and Ventral Prefrontal Cortices. *The Journal of Neuroscience*. 2003;23(21):7931-7939. <https://doi.org/10.1523/JNEUROSCI.23-21-07931.2003>.
- O'Doherty JP, Dolan RJ. (2006). The role of human orbitofrontal cortex in reward prediction and behavioral choice: Insights from neuroimaging. In: Zald DH, Rauch SL, editors. *The orbitofrontal cortex*. Oxford, United Kingdom: Oxford University Press; 2006. pp. 265–284.
- Overskeid G. The Slave of the Passions: Experiencing Problems and Selecting Solutions. *Review of General Psychology*. 2000;4(3):284-309. <https://doi.org/10.1037/1089-2680.4.3.284>.
- Parasuraman R, et Caggiano D. (2005). Neural and genetic assays of human mental workload. In D.K., McBride, & D., Schmorow, editors. *Quantifying human information processing* (Chap. 4), Lexington Books.
- Perrey S, Thedon T, Rupp T. NIRS in ergonomics: Its application in industry for promotion of health and human performance at work. *Int J Ind Ergon*. 2010;40(2):185-189. <https://doi.org/10.1016/j.ergon.2008.11.002>.
- 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(1):43-63. <https://doi.org/10.1006/cogp.1997.0659>.

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(2):609618. <https://doi.org/10.1016/j.neuroimage.2005.02.009>.

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(11):18711880. <https://doi.org/10.1111/j.1528-1167.2008.01711.x>.

Rothkirch M, Schmack K, Schlagenhauf F, Sterzer P. Implicit motivational value and salience are processed in distinct areas of orbitofrontal cortex. *NeuroImage*. 2012;62(3):17171725. <https://doi.org/10.1016/j.neuroimage.2012.06.016>.

Rygula R, Walker SC, Clarke HF, Robbins TW, Roberts AC. Differential Contributions of the Primate Ventrolateral Prefrontal and Orbitofrontal Cortex to Serial Reversal Learning. *The Journal of Neuroscience*. 2010;30(43):1455214559. <https://doi.org/10.1523/JNEUROSCI.2631-10.2010>.

Schmidt R, Leventhal DK, Mallet N, Chen F, Berke JD. Canceling actions involves a race between basal ganglia pathways. *Nature neuroscience*. 2013;16(8):11181124. <https://doi.org/10.1038/nn.3456>.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *The Journal of Neuroscience*. 2007;27(9):23492356. <https://doi.org/10.1523/JNEUROSCI.5587-06.2007>.

Smielewski P, Kirkpatrick P, Minhas P, Pickard JD, Czosnyka M. Can cerebrovascular reactivity be measured with near-infrared spectroscopy? *Stroke*. 1995;26(12):22852292. <https://doi.org/10.1161/01.str.26.12.2285>.

Sobotka SS, Davidson RJ, Senulis JA. Anterior brain electrical asymmetries in response to reward and punishment. *Electroencephalogr Clin Neurophysiol*. 1992;83(4):236247. [https://doi.org/10.1016/0013-4694\(92\)90117-Z](https://doi.org/10.1016/0013-4694(92)90117-Z).

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;223:143164. <https://doi.org/10.1016/bs.pbr.2015.07.024>.

Spielberg JM, Miller GA, Warren SL, Engels AS, Crocker LD, Banich MT, Sutton BP, Heller W. A brain network instantiating approach and avoidance motivation. *Psychophysiology*. 2012;49(9):12001214. <https://doi.org/10.1111/j.1469-8986.2012.01443.x>.

Swann NC, Cai W, Conner CR, Pieters TA, Claffey MP, George JS, Aron AR, Tandon N. 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(3):28602870. <https://doi.org/10.1016/j.neuroimage.2011.09.049>.

- Swick D, Ashley V, Turken AU. Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience*. 2008;9:102. <https://doi.org/10.1186/1471-2202-9-102>.
- Szatkowska I, Bogorodzki P, Wolak T, Marchewka A, Szeszkowski W. The effect of motivation on working memory: An fMRI and SEM study. *Neurobiol Learn Mem*. 2008;90(2):475478. <https://doi.org/10.1016/j.nlm.2008.06.001>.
- Tops M, Boksem MAS, Quirin M, IJzerman H, Koole SL. (2014). Internally directed cognition and mindfulness: An integrative perspective derived from predictive and reactive control systems theory. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00429>.
- 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(4):649672. <https://doi.org/10.1037/a0013170>.
- 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 Economics*. 2014;7(4):203218. <https://doi.org/10.1037/npe0000026>.
- Wager TD, Smith EE. Neuroimaging studies of working memory: A meta-analysis. *Cognitive Affective Behavioral Neuroscience*. 2003;3(4):255274. <https://doi.org/10.3758/cabn.3.4.255>.
- Wallis JD. Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci*. 2007;30:3156. <https://doi.org/10.1146/annurev.neuro.30.051606.094334>.
- Weinberger DR. A connectionist approach to the prefrontal cortex. *The Journal of Neuropsychiatry Clinical Neurosciences*. 1993;5(3):241253. <https://doi.org/10.1176/jnp.5.3.241>.
- Wilcox T, Biondi M. FNIRS in the developmental sciences. *Wiley Interdiscip Rev Cogn Sci*. 2015;6(3):263283. <https://doi.org/10.1002/wcs.1343>.
- 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(5):10191027. <https://doi.org/10.1093/cercor/bhn147>.
- 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(12):e82169. <https://doi.org/10.1371/journal.pone.0082169>.

Figures

Figure 1

Reversal-learning task. (A) Negative pair: antelope was the incorrect (-100 \$) and cow the correct (+0 \$) answer. (B) 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 (C) and duck became the incorrect answer (gains: 0 \$) and butterfly the correct answer (gain: +100 \$) for the positive pair (D). 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).

Figure 2

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.

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, is the average of eight participants' answers along the four recordings blocks so at lowest 32 acquisitions. The perseverative errors (C) indicate the eight participants' average number of successive incorrect answers after the reversal along the four recordings blocks; a lower score indicates that participants understood the reversal very quickly and applied it at the second repetition.

Figure 4

(A)

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, of which the 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 t05-15 are shown. Regions shaded in blue indicate under-activation, whereas shaded in red indicate strong activation. Shaded error bars indicate 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 statistics t-test p-value 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 statistics t-test p-value are shown below for the reward condition relative to the punishment condition. (D) Hemodynamic activity for 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 statistics t-test p-value are shown below for the reward condition relative to the punishment condition. (E) Hemodynamic activity for 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 statistics t-test p-value 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 statistics t-test p-value are shown below for the reward condition relative to the punishment condition.

Figure 5
(A)

(C)

	Reward	Punishment	Reward vs Punishment	Post Hoc
Amplitude	l-DLPFC	Significant (HbO)	Significant (HbO & HbT)	Significant vs r-DLPFC under Reward (HbO)
	l-IFG	Significant (HbO)	Tendency (HbO)	Tendency (HbO)
	r-DLPFC	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)	Tendency (HbO)
	r-DLPFC	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)

Figure 5

(A) Fisher LSD post hoc analysis for HbO for t05-15. 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 . 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.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1rvwd1.pdf](#)