

Neural interactions in the human frontal cortex dissociate reward and punishment learning

Reviewed Preprint

v2 • June 13, 2024

Revised by authors

Reviewed Preprint

v1 • November 9, 2023

Etienne Combrisson , Ruggero Basanisi, Maëlle C. M. Gueguen, Sylvain Rheims, Philippe Kahane, Julien Bastin, Andrea Brovelli 

Institut de Neurosciences de la Timone, Aix Marseille Université, UMR 7289 CNRS, 13005, Marseille, France • Univ. Grenoble Alpes, Inserm, U1216, Grenoble Institut Neurosciences, GIN, Grenoble, France • Department of Functional Neurology and Epileptology, Hospices Civils de Lyon and University of Lyon, Lyon, France • Univ. Grenoble Alpes, Inserm, U1216, CHU Grenoble Alpes, Grenoble Institut Neurosciences, GIN, Grenoble, France

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

How human prefrontal and insular regions interact while maximizing rewards and minimizing punishments is unknown. Capitalizing on human intracranial recordings, we demonstrate that the functional specificity toward reward or punishment learning is better disentangled by interactions compared to local representations. Prefrontal and insular cortices display non-selective neural populations to rewards and punishments. Non-selective responses, however, give rise to context-specific interareal interactions. We identify a reward subsystem with redundant interactions between the orbitofrontal and ventromedial prefrontal cortices, with a driving role of the latter. In addition, we find a punishment subsystem with redundant interactions between the insular and dorsolateral cortices, with a driving role of the insula. Finally, switching between reward and punishment learning is mediated by synergistic interactions between the two subsystems. These results provide a unifying explanation of distributed cortical representations and interactions supporting reward and punishment learning.

eLife assessment

This is an **important** information-theoretic re-analysis of human intracranial recordings during reward and punishment learning. It provides **convincing** evidence that reward and punishment learning is represented in overlapping regions of the brain while relying on specific inter-regional interactions. This preprint will be interesting to researchers in systems and cognitive neuroscience.

<https://doi.org/10.7554/eLife.92938.2.sa3>

Introduction

Reward and punishment learning are two key facets of human and animal behavior, because they grant successful adaptation to changes in the environment and avoidance of potential harm. These learning abilities are formalized by the law of effect (Thorndike, 1898; Bouton, 2007) and they pertain the goal-directed system, which supports the acquisition of action-outcome contingencies and the selection of actions according to expected outcomes, as well as current goal and motivational state (Dickinson and Balleine, 1994; Balleine and Dickinson, 1998; Balleine and O'Doherty, 2010; Dolan and Dayan, 2013; Balleine, 2019).

At the neural level, a first hypothesis suggests that these abilities are supported by distinct frontal areas (Pessiglione and Delgado, 2015; Palminteri and Pessiglione, 2017). Indeed, an anatomical dissociation between neural correlates of reward and punishment prediction error (PE) signals has been observed. PE signals are formalized by associative models (Rescorla et al., 1972) and reinforcement learning theory (Sutton and Barto, 2018) as the difference between actual and expected action's outcomes. Reward prediction error (RPE) signals have been observed in the midbrain, ventral striatum and ventromedial prefrontal cortex (vmPFC) (Schultz et al., 1997; O'Doherty et al., 2001, 2004; Pessiglione et al., 2006; D'Ardenne et al., 2008; Steinberg et al., 2013; Palminteri et al., 2015; Gueguen et al., 2021). Punishment prediction error (PPE) signals have been found in the anterior insula (aINS), dorsolateral prefrontal cortex (dlPFC), lateral orbitofrontal cortex (lOFC) and amygdala (O'Doherty et al., 2001; Seymour et al., 2005; Pessiglione et al., 2006; Yacubian et al., 2006; Gueguen et al., 2021). Evidence from pharmacological manipulations and lesion studies also indicates that reward and punishment learning can be selectively affected (Frank et al., 2004; Bódi et al., 2009; Palminteri et al., 2009, 2012). Complementary evidence, however, suggests that reward and punishment learning may instead share common neural substrates. Indeed, hubs of the reward circuit, such as the midbrain dopamine systems and vmPFC, contain neural populations encoding also punishments (Tom et al., 2007; Matsumoto and Hikosaka, 2009; Plassmann et al., 2010; Monosov and Hikosaka, 2012). Taken together, it is still unclear whether reward and punishment learning recruit complementary cortical circuits and whether differential interactions between frontal regions support the encoding of RPE and PPE.

To address this issue, we repose on recent literature proposing that learning reflects a network phenomenon emerging from neural interactions distributed over cortical-subcortical circuits (Bassett and Mattar, 2017; Hunt and Hayden, 2017; Averbeck and Murray, 2020; Averbeck and O'Doherty, 2022). Indeed, cognitive functions emerge from the dynamic coordination over large-scale and hierarchically organized networks (Varela et al., 2001; Bressler and Menon, 2010; Reid et al., 2019; Panzeri et al., 2022; Thiebaut de Schotten and Forkel, 2022; Miller et al., 2024; Noble et al., 2024) and accumulating evidence supports that information about task variables is widely distributed across brain circuits, rather than anatomically localized (Parras et al., 2017; Saleem et al., 2018; Steinmetz et al., 2019; Urai et al., 2022; Voitov and Mrcic-Flogel, 2022).

Accordingly, we investigated whether reward and punishment learning arise from complementary cortico-cortical functional interactions, defined as statistical relationships between the activity of different cortical regions (Panzeri et al., 2022), within and/or between brain regions of the frontal cortex. In particular, we investigated whether reward and punishment prediction errors are encoded by redundancy- and/or synergy-dominated functional interactions in the frontal cortex. The search for synergy- and redundancy-dominated interactions is motivated by recent hypotheses suggesting that a trade-off between redundancy for robust sensory and motor functions and synergistic interaction may be important for flexible higher cognition (Luppi et al., 2024). On one hand, we reasoned that redundancy-dominated brain networks may be

associated with neural interactions subserving similar functions. Redundant interactions may appear in collective states dominated by oscillatory synchronization (Engel et al., 2001 [↗](#); Varela et al., 2001 [↗](#); Buzsáki and Draguhn, 2004 [↗](#); Fries, 2015 [↗](#)) or resonance phenomena (Vinck et al., 2023 [↗](#)). Such collective states may give rise to selective patterns of information flow (Buehlmann and Deco, 2010 [↗](#); Kirst et al., 2016 [↗](#); Battaglia and Brovelli, 2020 [↗](#)). On the other, synergy-dominated brain networks may be associated with functionally-complementary interactions. Indeed, synergistic interactions have reported between distant transmodal regions during high-level cognition (Luppi et al., 2022 [↗](#)) and, at the microscale, in populations of neurons within a cortical column of the visual cortex and across areas of the visuomotor network (Nigam et al., 2019 [↗](#); Varley et al., 2023 [↗](#)). The notion of redundant and synergistic interactions resonates with the hypothesis that brain interactions regulate segregation and integration processes to support cognitive functions (Wang et al. 2021; Deco et al. 2015; Sporns 2013; Finc et al. 2020; Cohen and D'Esposito 2016; Braun et al. 2015; Shine et al. 2016).

In order to study redundancy- and synergy-dominated interactions, we used formal definitions from Partial Information Decomposition (PID; Williams and Beer, 2010 [↗](#); Wibral et al., 2017 [↗](#); Lizier et al., 2018 [↗](#)). The PID decomposes the total information that a set of source variables (i.e., pairs of brain signals) encodes about a specific target variable (i.e., prediction errors) into components representing shared (redundant) encoding between the variables, unique encoding by some of the variables, or synergistic encoding in the combination of different variables. Within this framework, we used a metric known as interaction information (McGill, 1954 [↗](#); Ince et al., 2017 [↗](#)), which quantifies whether a three-variable interaction (i.e., pairs of brain regions and the PE variable) is either synergy- and redundancy-dominated. We predicted that redundancy-dominated functional interactions would engage areas with similar functional properties (e.g., those encoding RPE), whereas synergy-dominated relations would be observed between areas performing complementary functions (e.g., the encoding of RPE and PPE).

We investigated neural interactions within and between four cortical regions, namely the aINS, dlPFC, IOFC and vmPFC, by means of intracerebral EEG (iEEG) data collected from epileptic patients while performing a reinforcement learning task (Gueguen et al., 2021 [↗](#)). We found various proportions of intracranial recordings encoding uniquely RPE or PPE signals or both, suggesting a local mixed representation of PEs. We then identified two distinct learning-related subsystems dominated by redundant interactions. A first subsystem with RPE-only interactions between the vmPFC and IOFC, and a second subsystem with PPE-only interactions between the aINS and dlPFC. Within each redundant-dominated subsystem, we demonstrated differential patterns of directional interactions, with the vmPFC and aINS playing a driving role in the reward and punishment learning circuits, respectively. Finally, these two subsystems interacted during the encoding of PE signals irrespectively of the context (reward or punishment), through synergistic collaboration between the dlPFC and vmPFC. We concluded that the functional specificity toward reward or punishment learning is better disentangled by interactions compared to local representations. Overall, our results provide a unifying explanation of distributed cortical representations and interactions supporting reward and punishment learning.

Results

iEEG data, behavioral task and computational modeling

We analyzed iEEG data from sixteen pharmacoresistant epileptic patients implanted with intracranial electrodes (Gueguen et al., 2021 [↗](#)). A total of 248 iEEG bipolar derivations located in the aINS, dlPFC, vmPFC and IOFC regions (Fig. 1A [↗](#)) and 1788 pairs of iEEG signals, both within and across brain regions (Fig. 1B [↗](#)) were selected for further analysis. Single subject anatomical repartition is shown in Fig. S1 [↗](#). Participants performed a probabilistic instrumental learning task and had to choose between two cues to either maximize monetary gains (for reward cues) or

minimize monetary losses (for punishment cues) (**Fig. 1C**). Overall, they selected more monetary gains and avoided monetary losses but the task structure was designed so that the number of trials was balanced between reward and punishment conditions (**Fig. 1D**).

We estimated trial-wise prediction errors by fitting a Q-learning model to behavioral data. Fitting the model consisted in adjusting the constant parameters to maximize the likelihood of observed choices. We used three constant parameters: i) the learning rate α accounting for how fast participants learned new pairs of cues; ii) the choice temperature β to model different levels of exploration and exploitation; iii) Θ parameter to account for the tendency to repeat the choice made in the previous trial. The RPE and PPE were obtained by taking the PE for rewarding and punishing pairs of cues, respectively. RPE and PPE showed high absolute values early during learning and tended toward zero as participants learned to predict the outcome (**Fig. 1E**). The convergence toward zero of RPE and PPE was stable at the single subject level (**Fig. S2**).

Local mixed encoding of PE signals

At the neural level, we first investigated local correlates of prediction error signals by studying whether RPEs and PPEs are differentially encoded in prefrontal and insular regions. To this end, we performed model-based information theoretical analyses of iEEG gamma activities by computing the mutual information (MI) between the across-trials modulations in RPE or PPE signals and the gamma band power in the aINS, dlPFC, IOFC and vmPFC. The MI allowed us to detect both linear and non-linear relationships between the gamma activity and the PE. Preliminary spectrally-resolved analyses showed that the frequency range significantly encoding prediction errors was between 50 and 100 Hz (**Fig. S3**). We thus extracted for each trial time-resolved gamma power within the 50 to 100 Hz range using a multi-taper approach for further analyses. MI analysis between gamma power and prediction error signals displayed significant group-level effects in all four cortical regions (**Fig. 2A**) and globally reproduced previous findings based on general linear model analyses (Gueguen et al., 2021). Interestingly, we observed a clear spatial dissociation between reward and punishment PE signaling. Whereas the vmPFC and dlPFC displayed complementary functional preferences for RPE and PPE, respectively, the aINS and the IOFC carried similar amounts of information about both R/PPE (**Figure 2A**).

To better characterize the spatial granularity of PE encoding, we further studied the specificity of individual brain regions by categorizing bipolar derivations as either: i) RPE-specific; ii) PPE-specific; iii) PE-unspecific responding to both R/PPE; iv) PE-irrelevant (i.e. non-significant ones) (**Fig. 2B**). All regions displayed a local mixed encoding of prediction errors with temporal dynamics peaking around 500 ms after outcome presentation. The vmPFC and dlPFC differentially responded to reward and punishment PEs, and contained approximately 30% of RPE- and PPE-specific contacts, respectively. In both regions, the proportion of RPE- and PPE-specific bipolar derivations was elevated for approximately 1s after outcome presentation. The IOFC also contained a large proportion of PPE-specific bipolar derivations, but displayed more transient dynamics lasting approximately 0.5s. The aINS had similar proportions of bipolar derivations specific for the RPE and PPE (20%), with temporal dynamics lasting approximately 0.75s. Importantly, all regions contained approximately 10% of PE-unspecific bipolar derivations that responded to both RPE and PPE, especially in the aINS and dlPFC. The remaining bipolar derivations were categorized as PE-irrelevant. A complementary analysis, conducted to evaluate inter-subject reproducibility, revealed that local encoding in the IOFC and vmPFC was represented in 30 to 50% of the subjects. In contrast, this encoding was found in 50 to 100% of the subjects in the aINS and dlPFC (**Fig. S4**).

Taken together, our results demonstrate that reward and avoidance learning are not supported by highly selective brain activations, but rather from a mixed or mixed encoding of RPE or PPE signals distributed over the prefrontal and insular cortices. Nevertheless, such distributed encoding seems to involve two complementary systems primarily centered over the vmPC and dlPFC, respectively.

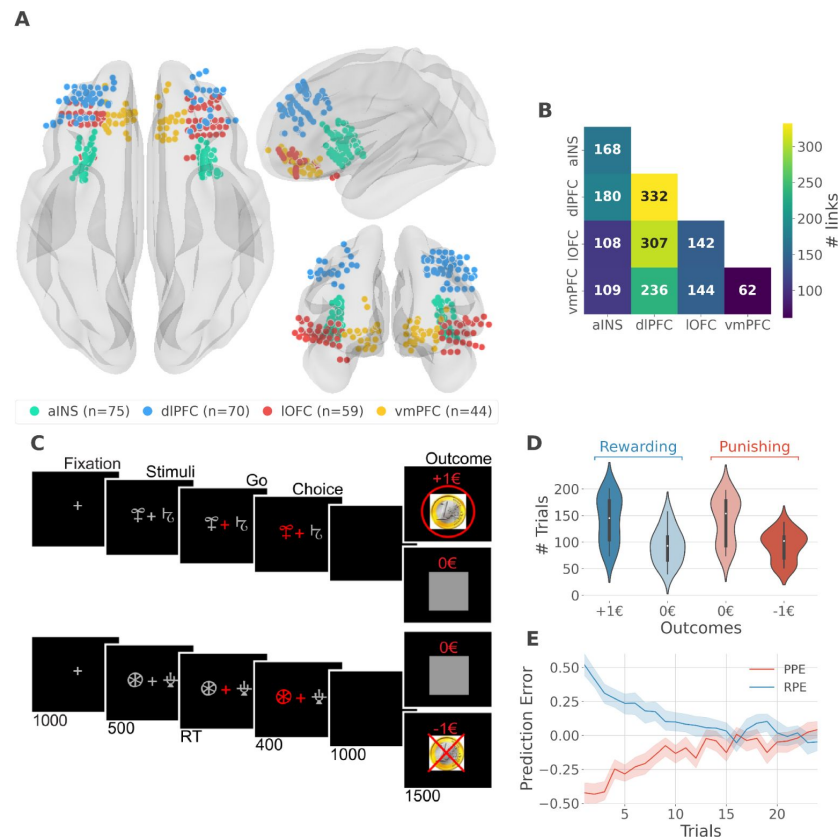


Figure 1.

iEEG implantation, behavioral task and computational modeling

(A) Anatomical location of intracerebral electrodes across the 16 epileptic patients. Anterior insula (aINS, $n=75$), dorsolateral prefrontal cortex (dlPFC, $n=70$), lateral orbitofrontal cortex (IOFC, $n=59$), ventromedial prefrontal cortex (vmPFC, $n=44$), (B) Number of pairwise connectivity links (i.e. within patients) within and across regions, (C) Example of a typical trial in the reward (top) and punishment (bottom) conditions. Participants had to select one abstract visual cue among the two presented on each side of a central visual fixation cross and subsequently observed the outcome. Duration is given in milliseconds, (D) Number of trials where participants received outcomes +1€ (142 ± 44 , mean $\pm$ std) vs. 0€ (93 ± 33) in the rewarding condition (blue) and outcomes 0€ (141 ± 42) vs. -1€ (93 ± 27) in the punishment condition (red), (E) Across participants trial-wise reward PE (RPE - blue) and punishment PE (PPE - red), $\pm$ 95% confidence interval.

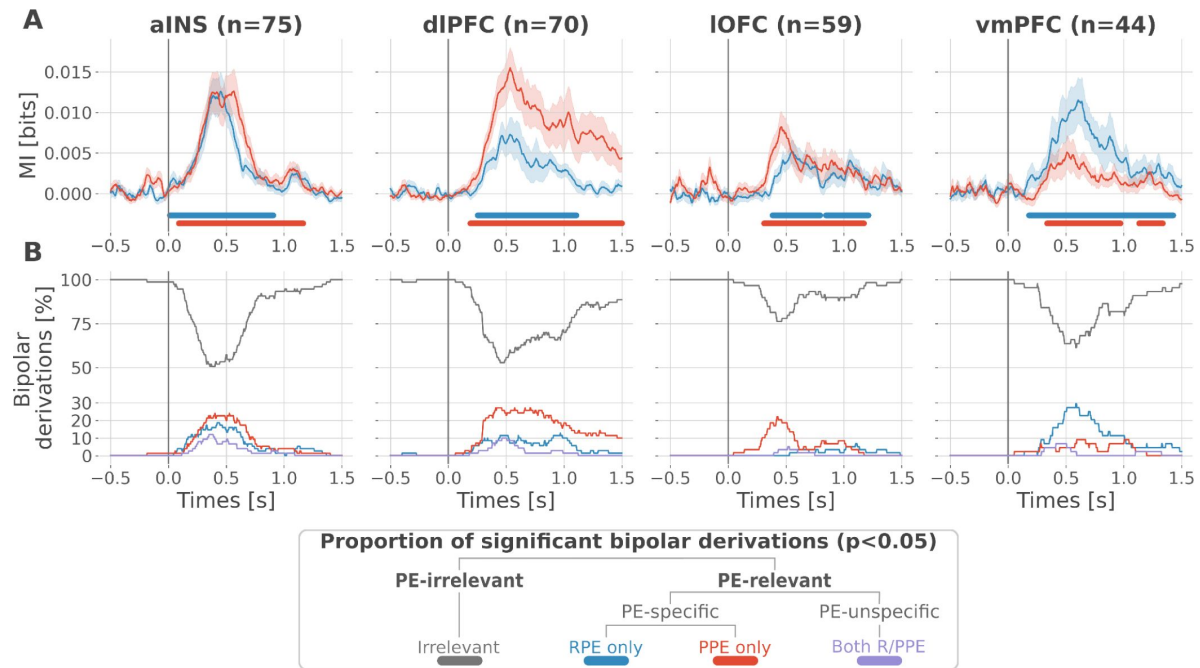


Figure 2.

Local mixed encoding of reward and punishment prediction error signals

(A) Time-courses of mutual information (MI in bits) estimated between the gamma power and the reward (blue) and punishment (red) PE signals. The solid line and the shaded area represent the mean and SEM of the across-contacts MI. Significant clusters of MI at the group level are plotted with horizontal bold lines ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs), (B) Instantaneous proportions of task-irrelevant (gray) and task-relevant bipolar derivations presenting a significant relation with either the RPE (blue), the PPE (red) or with both RPE and PPE (purple). Data is aligned to the outcome presentation (vertical line at 0 seconds).

Encoding of PE signals occur with redundancy-dominated subsystems

To better understand the observed complex encoding of reward and punishment PEs, we tested the hypothesis that functional dissociations occur with differential and distributed interactions between prefrontal and insular cortices. To address this question, we performed model-based network-level analyses based on the Partial Information Decomposition (PID) framework (Williams and Beer, 2010; Wibral et al., 2017; Lizier et al., 2018). We particularly used the interaction information (McGill, 1954; Ince et al., 2017) to quantify whether a three-variable interaction (i.e., pairs of brain regions and the PE variable) is either synergy- and redundancy-dominated (Williams and Beer, 2010). Indeed, interaction information (II) can be either positive or negative. A negative value indicates a net redundancy (i.e., pair of recordings are carrying similar information about the PE), whereas a positive value indicates a net synergistic effect (i.e., pair of recordings are carrying complementary information about the PE). We computed the time-resolved II across trials between the gamma activity of pairs of iEEG signals and PEs. To differentiate cortico-cortical interactions for reward and punishment learning, we first calculated the II separately for RPEs and PPEs. RPE- and PPE-specific analyses exclusively showed negative modulations of II, therefore indicating the presence of redundancy-dominated local and long-range interactions (Fig. 3).

To better characterize the local interactions encoding reward and punishment PEs, we computed the II between pairs of gamma band signals recorded within the aINS, dlPFC, IOFC and vmPFC. Within-region II analyses showed that significant RPE-specific interactions were exclusively observed in the vmPFC and IOFC, whereas PPE-specific interactions were present only in the dlPFC. In addition, the aINS was found to display both RPE- and PPE-specific interactions (Fig. 3A). A relevant sign of high specificity for either reward or punishment PE signals was the presence of a significant cluster dissociating RPE and PPE in the vmPFC and dlPFC only (green clusters in Fig. 3A).

To investigate the nature of long-range interactions, we next computed the II for RPE and PPE between signals from different brain regions (Fig. 3B). Similarly, results exclusively showed redundancy-dominated interactions (i.e., negative modulations). RPE-specific interactions were observed between the IOFC and vmPFC, whereas PPE-specific interactions were observed between the aINS and dlPFC and to a smaller extent between the dlPFC and IOFC, peaking 500 ms after outcome presentation. A significant difference between RPE and PPE was exclusively observed in the IOFC-vmPFC and aINS-dlPFC interactions, but not between dlPFC and IOFC (green clusters in Fig. 3B). The analysis of inter-subject reproducibility revealed that both within-area and across-area significant redundant interactions were carried by 30 to 60% of the subjects (Fig. S5). Taken together, we conclude that the encoding of RPE and PPE signals occur with redundancy-dominated subsystems that differentially engage prefronto-insular regions.

Contextual directional interactions within redundant subsystems

Previous analyses of II are blind to the direction of information flows. To address this issue, we estimated the transfer entropy (TE) (Schreiber, 2000) on the gamma power during the rewarding (TE_{Rew}) and punishment conditions (TE_{Pun}), between all possible pairs of contacts. As a reminder, the TE is an information-theoretic measure that quantifies the degree of directed statistical dependence or “information flow” between time series, as defined by the Wiener-Granger principle (Wiener, 1956; Granger, 1969). Delay-specific analyses of TE showed that a maximum delay of information transfer between pairs of signals comprised an interval between 116 and 236 msec (Fig. S6). We thus computed the TE for all pairs of brain regions within this range of delays and detected temporal clusters where the TE significantly differed between conditions ($TE_{\text{Rew}} > TE_{\text{Pun}}$ or $TE_{\text{Pun}} > TE_{\text{Rew}}$). Only two pairs of brain regions displayed statistically-significant modulations in TE (Fig. 4). We observed that the TE from the aINS to the

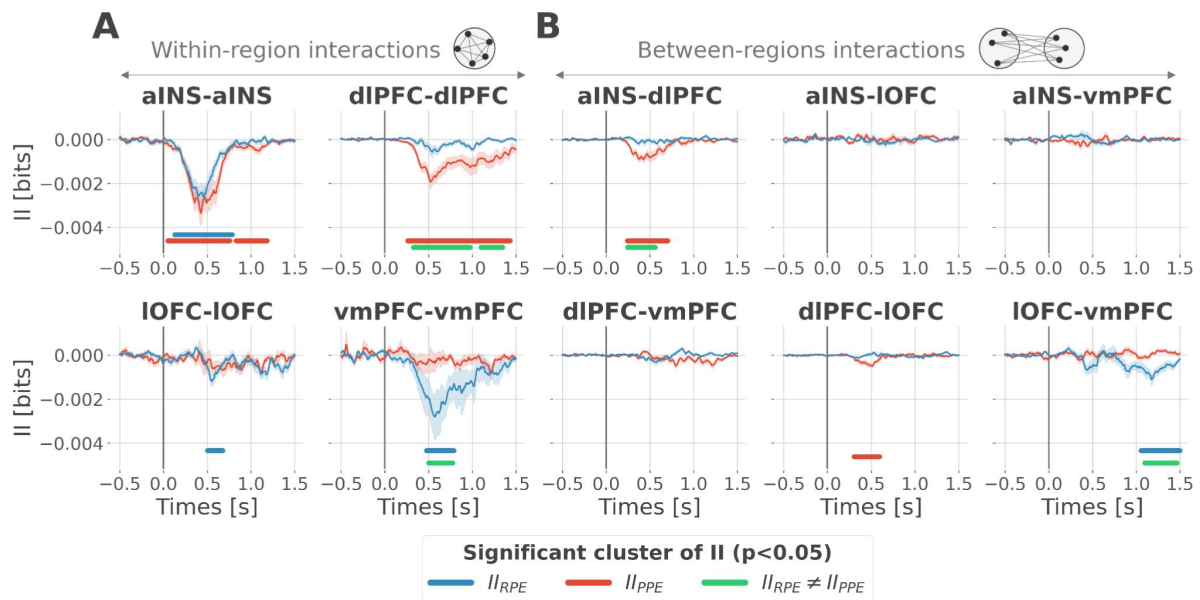


Figure 3.

Encoding of PE signals occurs with redundancy-dominated subsystems

Dynamic interaction information (II in bits) within-(A) and between-regions (B) about the RPE (II_{RPE}) and PPE (II_{PPE}) are plotted in blue and red. Significant clusters of II_{RPE} and II_{PPE} are displayed with horizontal bold blue and red lines ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs). Significant differences between II_{RPE} and II_{PPE} are displayed in green. Shaded areas represent the SEM. The vertical gray line at 0 seconds represents the outcome presentation.

dlPFC ($TE_{aINS \rightarrow dlPFC}$) peaked at approximately 400 msec after outcome onset and was significantly stronger during the punishment condition compared to the rewarding condition. By contrast, the information flow around ~800ms from the vmPFC to the IOFC ($TE_{vmPFC \rightarrow IOFC}$) was significantly stronger during the rewarding condition. No other brain interactions were found significant (**Fig. S7**). Overall, these results demonstrate that the two redundancy-dominated RPE- and PPE-specific networks (**Fig. 3B**) are characterized by differential directional interactions. The vmPFC and aINS act as drivers in the two systems, whereas the dlPFC and IOFC play the role of receivers, thus suggesting a flow of PE-specific information within the network.

Integration of PE signals occurs with synergy-dominated interactions between segregated sub-systems

Since learning required participants to concurrently explore rewarding and punishment outcomes, we finally investigated the nature of cortico-cortical interactions encoding both RPE and PPE signals. We estimated the II about the full PEs, i.e. the information carried by co-modulation of gamma power between all pairs of contacts about PE signals (Fig S8). Encoding of PEs was specifically associated with significantly positive II between the dlPFC and vmPFC ($II_{dlPFC-vmPFC}$; **Fig. 5A**). Such between-regions synergy-dominated interaction occurred approximately between 250 and 600 ms after outcome onset.

We then investigated if the synergy between the dlPFC and vmPFC encoding global PEs could be explained by their respective local specificity. Indeed, we previously reported larger proportions of recordings encoding the PPE in the dlPFC and the RPE in the vmPFC (**Fig. 2B**). Therefore, it is possible that the positive $II_{dlPFC-vmPFC}$ could be mainly due to complementary roles where the dlPFC brings information about the PPE only and the vmPFC brings information to the RPE only. To test this possibility, we computed the $II_{dlPFC-vmPFC}$ for groups of bipolar derivations with different local specificity. As a reminder, bipolar derivations were previously categorized as RPE or PPE specific if their gamma activity were modulated according to the RPE only, to the PPE only or to both (**Fig. 2B**). We obtained four categories of II. The first two categories, named $II_{RPE-RPE}$ and $II_{PPE-PPE}$, reflect the II estimated between RPE- and PPE-bipolar derivations from the dlPFC and vmPFC. The third category ($II_{PPE-RPE}$) refers to the II estimated between PPE-specific bipolar recordings from the dlPFC and RPE-specific bipolar recordings from the vmPFC. Finally, the fourth category, named II_{Mixed} , includes the remaining possibilities (i.e., RPE-Both, PPE-Both and Both-Both) (**Fig. S9**). Interestingly, we found significant synergistic interactions between recordings with mixed specificity i.e. $II_{PPE-RPE}$ and II_{Mixed} between 250 and 600 ms after outcome onset (**Fig. 5B** and **5C**). Consequently, the $II_{dlPFC-vmPFC}$ is partly explained by the dlPFC and vmPFC carrying PPE- and RPE-specific information ($II_{PPE-RPE}$) together with interactions between non-specific recordings (II_{Mixed}). In addition, we simulated data to demonstrate that synergistic interactions can emerge between regions with the same local specificity (**Fig. S10**). Taken together, the integration of the global PE signals occurred with synergistic interaction between recordings with mixed specificity from the dlPFC and vmPFC.

Discussion

Our study revealed the presence of specific functional interactions between prefrontal and insular cortices about reward and punishment prediction error signals. We first provided evidence for a mixed encoding of reward and punishment prediction error signals in each cortical region. We then identified a first subsystem specifically encoding RPEs with emerging redundancy-dominated interactions within and between the vmPFC and IOFC, with a driving role of the vmPFC. A second subsystem specifically encoding PPEs occurred with redundancy-dominated interactions within

Figure 4.

Contextual modulation of information transfer

Time courses of transfer entropy (TE, in bits) from the aINS to the dIPFC (aINS → dIPFC) and from the vmPFC to the IOFC (vmPFC → IOFC), estimated during the rewarding condition (TE_{Rew} in blue) and punishing condition (TE_{Pun} in red). Significant differences ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs) of TE between conditions are displayed with horizontal bold lines (blue for $TE_{\text{Rew}} > TE_{\text{Pun}}$ and red for $TE_{\text{Pun}} > TE_{\text{Rew}}$). Shaded areas represent the SEM. The vertical gray line at 0 seconds represents the outcome presentation.

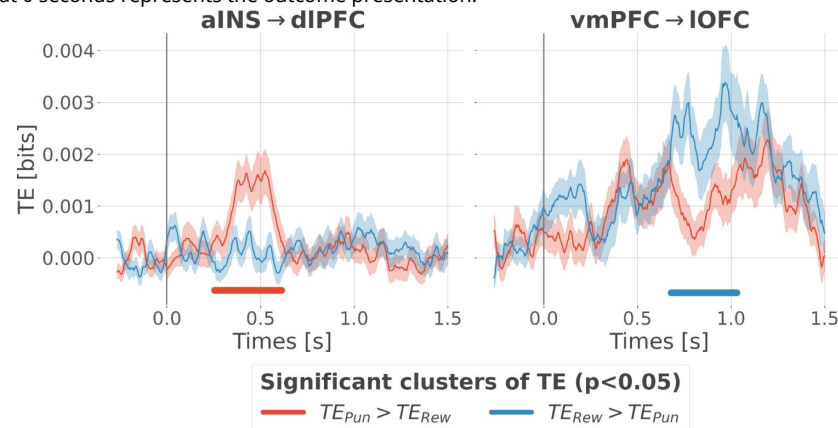
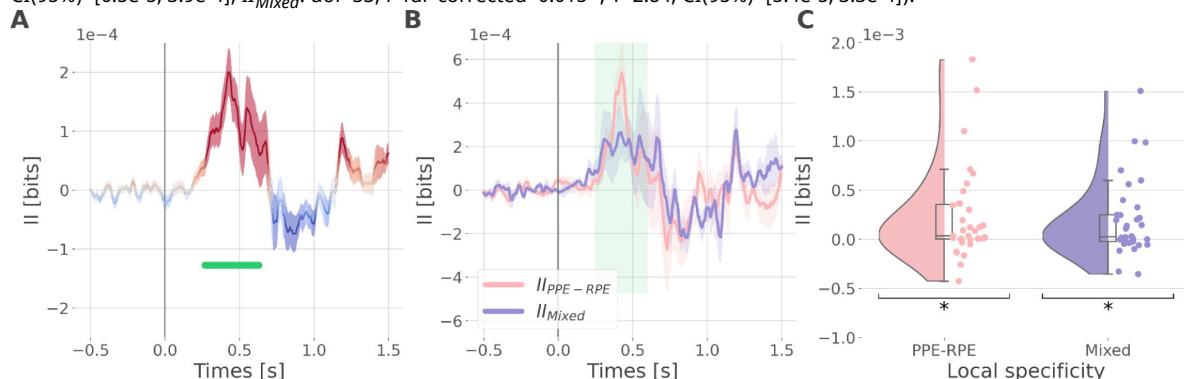


Figure 5.

Synergistic interactions about the full PE signals between recordings of the dIPFC and vmPFC

(A) Dynamic interaction information (II in bits) between the dIPFC and vmPFC about the full prediction error ($II_{\text{dIPFC-vmPFC}}$). Hot and cold colors indicate synergy- and redundancy-dominated II about the full PE. Significant clusters of II are displayed with a horizontal bold green line ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs). Shaded areas represent the SEM. The vertical gray line at 0 seconds represents the outcome presentation. (B) Dynamic $II_{\text{dIPFC-vmPFC}}$ binned according to the local specificity PPE-RPE ($II_{\text{PPE-RPE}}$ in pink) or mixed (II_{Mixed} in purple) (C) Distributions of the mean of the $II_{\text{PPE-RPE}}$ and II_{Mixed} for each pair of recordings ($II_{\text{PPE-RPE}}$: one sample t-test against 0; dof=34; P fdr-corrected=0.015*; T=2.86; CI(95%)=[6.5e-5, 3.9e-4]; II_{Mixed} : dof=33; P fdr-corrected=0.015*; T=2.84; CI(95%)=[5.4e-5, 3.3e-4]).



and between the aINS and dlPFC, with a driving role of the aINS. Switching between the encoding of reward and punishment PEs involved a synergy-dominated interaction between these two systems mediated by interactions between the dlPFC and vmPFC (Fig. 6 [↗](#)).

Local mixed representations of prediction errors

Amongst the four investigated core-learning regions, the vmPFC was the only region to show a higher group-level preference for RPEs. This supports the notion that the vmPFC is functionally more specialized for the processing outcomes in reward learning, as previously put forward by human fMRI meta-analyses (Yacubian et al., 2006 [↗](#); Diekhof et al., 2012 [↗](#); Bartra et al., 2013 [↗](#); Garrison et al., 2013 [↗](#); Fouragnan et al., 2018 [↗](#)). The dlPFC, instead, showed a stronger selectivity for punishment PE, thus supporting results from fMRI studies showing selective activations for aversive outcomes (Liu et al., 2011 [↗](#); Garrison et al., 2013 [↗](#); Fouragnan et al., 2018 [↗](#)). On the contrary, the aINS and IOFC did not show clear selectivity for either reward or punishment PEs. The aINS carried a comparable amount of information about the RPE and PPE, thus suggesting that the insula is part of the surprise-encoding network (Fouragnan et al., 2018 [↗](#); Loued-Khenissi et al., 2020 [↗](#)). Previous study reported a stronger link between the gamma activity of the aINS and the PPE compared to the RPE (Gueguen et al., 2021 [↗](#)). This discrepancy in the results could be explained by the measures of information we are using here that are able to detect both linear and non-linear relationships between gamma activity and PE signals (Ince et al., 2017 [↗](#)). The IOFC showed an initial temporal selectivity for PPE followed by a delayed one about the RPE. This is in accordance with fMRI and human intracranial studies which revealed that the IOFC was activated when receiving punishing outcomes, but also contains reward-related information (O'Doherty et al., 2001 [↗](#); Saez et al., 2018 [↗](#); Gueguen et al., 2021 [↗](#)).

By taking advantage of the multi-site sampling of iEEG recordings, we quantified the heterogeneity in functional selectivity within each area and showed that the region-specific tendency toward either RPE or PPEs (Fig. 2A [↗](#)) could be explained by the largest domain-specific proportion of contacts (Fig. 2B [↗](#)). In other words, if a region showed a larger proportion of contacts being RPE-specific, the amount of information about the RPE at the group-level was also larger. Interestingly, we observed that 5 to 20% of contacts within a given region encoded both the RPE and PPE, thus revealing local mixed representations. Consequently, a strict dichotomous classification of learning-related areas as either reward and punishment may fail to capture important properties of the individual nodes of the learning circuit, such as the functional heterogeneity in the encoding of PEs. These results suggest that the human prefrontal cortex exhibits a mixed local selectivity for prediction error signals at the mesoscopic scale. This view is in line with recent literature showing that prefrontal cortex contains single neurons exhibiting mixed selectivity for multiple task variables (Meyers et al., 2008 [↗](#); Rigotti et al., 2013 [↗](#); Stokes et al., 2013 [↗](#); Panzeri et al., 2015 [↗](#); Parthasarathy et al., 2017 [↗](#); Bernardi et al., 2020 [↗](#)). In the learning domain, single-unit studies have reported neurons encoding both rewarding and aversive outcomes in the OFC of the primate (Morrison and Salzman, 2009 [↗](#); Monosov and Hikosaka, 2012 [↗](#); Hirokawa et al., 2019 [↗](#)). Mixed selectivity provides computational benefits, such as increasing the number of binary classifications, improving cognitive flexibility, and simplifying readout by downstream neurons (Fusi et al., 2016 [↗](#); Helfrich and Knight, 2019 [↗](#); Ohnuki et al., 2021 [↗](#); Panzeri et al., 2022 [↗](#)). We suggest that the encoding of cognitive variables such as prediction error signals is supported by similar principles based on mixed selectivity at the meso- and macroscopic level, and may provide a natural substrate for cognitive flexibility and goal-directed learning (Rigotti et al., 2013 [↗](#)).

Redundancy-dominated interactions segregate reward and punishment learning subsystems

We then tested whether the encoding of RPE and PPE signals could be supported by differential cortico-cortical interactions within and between frontal brain regions. To do so, we exploited the interaction information (II) (McGill, 1954 [↗](#); Ince et al., 2017 [↗](#)) to quantify whether the amount of

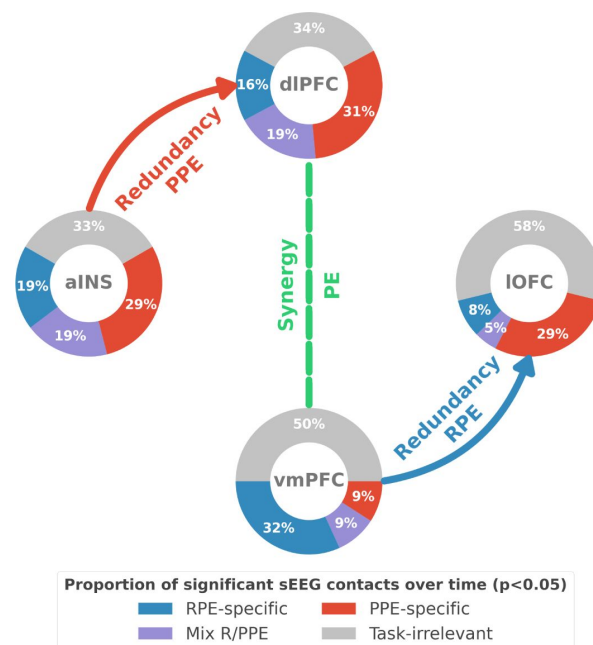


Figure 6.

Summary of findings

The four nodes represent the investigated regions, namely the anterior insula (aINS), the dorsolateral and ventromedial parts of the prefrontal cortex (dlPFC and vmPFC) and the lateral orbitofrontal cortex (IOFC). The outer disc represents the local mixed encoding i.e. the different proportions of contacts over time having a significant relationship between the gamma power and PE signals. In blue, is the proportion of contacts with a significant relation with the PE across rewarding trials (RPE-specific). Conversely, in red for punishment trials (PPE-specific). In purple, the proportion of contacts with a significant relationship with both the RPE and PPE. In gray, is the remaining proportion of non-significant contacts. Regarding interactions, we found that information transfer between aINS and dlPFC carried redundant information about PPE only and information transfer between vmPFC and IOFC about RPE only. This information transfer occurred with a leading role of the aINS in the punishment context and the vmPFC in the rewarding context. Finally, we found synergistic interactions between the dlPFC and the vmPFC about the full PE, without splitting into rewarding and punishing conditions.

information bound up in a pair of gamma responses and PE signals is dominated by redundant or synergistic interactions (Williams and Beer, 2010). The II revealed redundancy-dominated interactions specific for RPE and PPE in the vmPFC and the dlPFC, respectively (Fig. 3A). The aINS was the only region for which the between-contacts II did not increase the functional selectivity, with large redundant interactions for both RPE and PPE signals. This suggests that within-area redundant interactions can potentially amplify the functional specificity, despite the presence of local mixed selectivity (Fig. 2A). Such “winner-take-all” competition could be implemented by mutual inhibition mechanisms, which have been suggested to be essential in reward-guided choice (Hunt et al., 2012; Jocham et al., 2012; Strait et al., 2014; Hunt and Hayden, 2017).

Across-areas interaction information revealed two subsystems with redundancy-dominated interactions. A reward subsystem with RPE-specific interactions between the IOFC and vmPFC, and a punishment subsystem with PPE-specific interactions between the aINS and dlPFC (Fig. 3B). Although a significant modulation selective for RPE was also present in the interaction between dlPFC and IOFC peaking around 500 ms after outcome presentation, a significant difference between the encoding of RPE and PPE was exclusively observed in the IOFC-vmPFC and aINS-dlPFC interactions (green clusters in Fig. 3B). This result suggests that the observed functionally-distinct learning circuits for RPE and PPEs are associated with differential cortico-cortical interactions, rather than distinct local properties. More generally, our results suggest that redundancy-based network-level interactions are related with the functional specificity observed in neuroimaging and lesion studies (Pessiglione and Delgado, 2015; Palminteri and Pessiglione, 2017).

We then investigated differential communication patterns and directional relations within the two redundancy-dominated circuits (Kirst et al., 2016; Palmigiano et al., 2017). We identified significant information routing patterns, dissociating reward and punishment learning (Fig. 4). Within the reward subsystem, the vmPFC played a driving role toward the IOFC only during the rewarding condition. Conversely, within the punishment subsystem, the aINS played a driving role toward the dlPFC only during the punishment condition. These results support the notion that redundancy-dominated cognitive networks are associated with the occurrence of information-routing capabilities, where signals are communicated on top of collective reference states (Battaglia and Brovelli, 2020).

Here, we quantified directional relationships between regions using the transfer entropy (Schreiber, 2000), which is a functional connectivity measure based on the Granger-Wiener causality principle. Tract tracing studies in the macaque have revealed strong interconnections between the IOFC and vmPFC in the macaque (Carmichael and Price, 1996; Öngür and Price, 2000). In humans, cortico-cortical anatomical connections have mainly been investigated using diffusion magnetic resonance imaging (dMRI). Several studies found strong probabilities of structural connectivity between the anterior insula with the orbitofrontal cortex and dorsolateral part of the prefrontal cortex (Cloutman et al., 2012; Ghaziri et al., 2017), and between the IOFC and vmPFC (Heather Hsu et al., 2020). In addition, the statistical dependency (e.g. coherence) between the LFP of distant areas could be potentially explained by direct anatomical connections (Schneider et al., 2021; Vinck et al., 2023). Taken together, the existence of an information transfer might rely on both direct or indirect structural connectivity. However, here we also reported differences of TE between rewarding and punishing trials given the same backbone anatomical connectivity (Fig. 4). Our results are further supported by a recent study involving drug-resistant epileptic patients with resected insula who showed poorer performance than healthy controls in case of risky loss compared to risky gains (Von Siebenthal et al., 2017).

Encoding the full PE is supported by synergistic interactions between subsystems

Humans can flexibly switch between learning strategies that allow the acquisition of stimulus-action-outcomes associations in changing contexts. We investigated how RPE and PPE subsystems coordinated to allow such behavioral flexibility. To do so, we searched for neural correlates of PEs irrespectively of the context (reward or punishment learning) in between-regions interactions. We found that the encoding of global PE signals was associated with synergy-dominated interactions between the two subsystems, mediated by the interactions between the dlPFC and the vmPFC (**Fig. 5**). Importantly, such synergy-dominated interaction reveals that the joint representation of the dlPFC and vmPFC is greater than the sum of their individual contributions for the encoding of global PE signals. Thus, it suggests that successful adaptation in varying contexts requires both the vmPFC and dlPFC for the encoding of global PE signals.

Role of redundant and synergistic interactions in brain network coordination

At the macroscopic level, few studies investigated the potential role of redundant and synergistic interactions. By combining functional and diffusion MRI, recent work suggested that redundant interactions are predominantly associated with structurally coupled and functionally segregated processing. In contrast, synergistic interactions preferentially support functional integrative processes and complex cognition across higher-order brain networks (Luppi et al., 2022). Triadic synergistic interactions between the continuous spike counts recorded within and across areas of the visuomotor network have been shown to carry behaviorally-relevant information and to display the strongest modulations during the processing of visual information and movement execution (Varley et al., 2023). Finally, cortical representations of prediction error signals in the acoustic domain observed tone-related and instantaneous redundant interactions, such as time-lagged synergistic interactions within and across temporal and frontal regions of the auditory system (Gelens et al., 2023).

At the microscopic level, the amount of information encoded by a population of neurons can be modulated by pairwise and higher-order interactions, producing varying fractions of redundancy and synergy (Averbeck et al., 2006; Panzeri et al., 2015, 2022). Synergistic and redundant pairs of neurons can be identified by estimating the amount of information contained in the joint representation minus the sum of the information carried by individual neurons (Schneidman et al., 2003). Redundant coding is intricately linked to correlated activity (Gutnisky and Dragoi, 2008) and can spontaneously emerge due to the spatial correlations present in natural scenes by triggering neurons with overlapping receptive fields. Correlations between the trial-by-trial variations of neuronal responses could limit the amount of information encoded by a population (Ramon Bartolo et al., 2020; Kafashan et al., 2021) and facilitate readout by downstream neurons (Salinas and Sejnowski, 2001). While redundancy has been at the heart of heated debates and influential theories, such as efficient coding and redundancy compression in sensory areas (Barlow, 2001), synergy phenomena have been described to a lesser extent. Recently, a study reported synergistic coding in a V1 cortical column together with structured correlations between synergistic and redundant hubs (Nigam et al., 2019). Taken together, we suggest that population codes with balancing proportions of redundancy and synergy offer a good compromise between system robustness and resilience to cell loss and the creation of new information (Panzeri et al., 2022). We suggest that redundancy-dominated interactions confer robustness and network-level selectivity for complementary learning processes, which may lead to functional integration processes. On the other hand, synergy-dominated interactions seem to support neural interactions between redundancy-dominated networks, thus supporting functional integrative processes in the brain. In addition, our study suggests that redundant and synergistic interactions occur across multiple spatial scales from local to large-scale.

Conclusion

Our report of mixed representation of reward and punishment prediction error signals explains the discrepancy in the attribution of a functional specificity to the core learning cortical regions. Instead, we propose that functional specialization for reward and punishment PE signals occurs with redundancy-dominated interactions within the two subsystems formed by the vmPFC-IOFC and aINS-dlPFC, respectively. Within each subsystem, we observed asymmetric and directional interactions with the vmPFC and aINS playing a driving role in the reward and punishment learning circuits. Finally, switching between reward and punishment learning was supported by synergistic collaboration between subsystems. This supports the idea that higher-order integration between functionally-distinct subsystems are mediated by synergistic interactions. Taken together, our results provide a unifying view reconciling distributed cortical representations with interactions supporting reward and punishment learning. They highlight the relevance of considering learning as a network-level phenomenon by linking distributed and functionally redundant subnetworks through synergistic interactions hence supporting flexible cognition (Fedorenko and Thompson-Schill, 2014 [DOI](#); Petersen and Sporns, 2015 [DOI](#); Bassett and Mattar, 2017 [DOI](#); Hunt and Hayden, 2017 [DOI](#); Averbeck and Murray, 2020 [DOI](#); Averbeck and O'Doherty, 2022 [DOI](#)).

Methods

Data acquisition and experimental procedure

Intracranial EEG recordings

Intracranial electroencephalography (iEEG) recordings were collected from sixteen patients presenting pharmaco-resistant focal epilepsy and undergoing presurgical evaluation (33.5 ± 12.4 years old, 10 females). As the location of the epileptic foci could not be identified through noninvasive methods, neural activity was monitored using intracranial stereotactic electroencephalography. Multi-lead and semi-rigid depth electrodes were stereotactically implanted according to the suspected origin of seizures. The selection of implantation sites was based solely on clinical aspects. iEEG recordings were performed at the clinical neurophysiology epilepsy departments of Grenoble and Lyon Hospitals (France). iEEG electrodes had a diameter of 0.8mm, 2mm wide, 1.5mm apart and contained 8–18 contact leads (Dixi, Besançon, France). For each patient, 5 to 17 electrodes were implanted. Recordings were conducted using an audio-video-EEG monitoring system (Micromed, Treviso, Italy), which allowed simultaneous recording of depth iEEG channels sampled at 512 Hz (6 patients), or 1024 Hz (12 patients) [0.1–200 Hz bandwidth]. One of the contacts located in the white matter was used as a reference. Anatomical localizations of iEEG contacts were determined based on post-implant computed tomography scans or post-implant MRI scans coregistered with pre-implantation scans (Lachaux et al., 2003 [DOI](#); Mercier et al., 2022 [DOI](#)). All patients gave written informed consent and the study received approval from the ethics committee (CPP 09-CHUG-12, study 0907) and from a competent authority (ANSM no: 2009-A00239-48).

Limitations

iEEG have been collected from pharmacoresistant epileptic patients who underwent deep electrode probing for preoperative evaluation. However, we interpreted these data as if collected from healthy subjects and assumed that epileptic activity does not affect the neural realization of prediction error. To best address this question, we excluded electrodes contaminated with pathological activity and focused on task-related changes and multi-trial analysis to reduce the impact of incorrect or task-independent neural activations. Therefore, our results may benefit

from future replication in healthy controls using non-invasive recordings. Despite the aforementioned limitations, we believe that access to deep intracerebral EEG recordings of human subjects can provide privileged insight into the neural dynamics that regulate human cognition, with outstanding spatial, temporal, and spectral precision. In the long run, this type of data could help bridge the gap between neuroimaging studies and electrophysiological recordings in nonhuman primates.

Preprocessing of iEEG data

Bipolar derivations were computed between adjacent electrode contacts to diminish contributions of distant electric sources through volume conduction, reduce artifacts and increase the spatial specificity of the neural data. Bipolar iEEG signals can approximately be considered as originating from a cortical volume centered within two contacts (Brovelli et al., 2005 [↗](#); Bastin et al., 2016 [↗](#); Combrisson et al., 2017 [↗](#)), thus providing a spatial resolution of approximately 1.5-3mm (Lachaux et al., 2003 [↗](#); Jerbi et al., 2009 [↗](#); Mercier et al., 2022 [↗](#)). Recording sites with artifacts and pathological activity (e.g. epileptic spikes) were removed using visual inspection of all of the traces of each site and each participant.

Definition of anatomical regions of interest

Anatomical labeling of bipolar derivations was performed using the IntrAnat software (Deman et al., 2018 [↗](#)). The 3D T1 pre-implantation MRI gray/white matter was segmented and spatially normalized to obtain a series of cortical parcels using MarsAtlas (Auzias et al., 2016 [↗](#)) and the Destrieux atlas (Destrieux et al., 2010 [↗](#)). 3D coordinates of electrodes contacts were then coregistered on post-implantation images (MRI or CT). Each recording site (i.e., bipolar derivation) was labeled according to its position in a parcellation scheme in the participant's native space. Thus, the analyzed dataset only included electrodes identified to be in the gray matter. Four regions of interest (ROIs) were defined for further analysis: 1) the ventromedial prefrontal cortex (vmPFC) ROI was created by merging six (three per hemisphere) parcels in MarsAtlas (labeled PFCvm, OFCv and OFCvm in MarsAtlas) corresponding to the ventromedial prefrontal cortex and fronto-medial part of the orbitofrontal cortex, respectively; 2) the lateral orbitofrontal cortex (lOFC) ROI included four (two per hemisphere) MarsAtlas parcels (MarsAtlas labels: OFCvl and the OFCv); 3) the dorsolateral prefrontal cortex (dlPFC) ROI was defined as the inferior and superior bilateral dorsal prefrontal cortex (MarsAtlas labels: PFRdli and PFRdls); 4) the anterior insula (aINS) ROI was defined as the bilateral anterior part of the insula (Destrieux atlas labels: Short insular gyri, anterior circular insular sulcus and anterior portion of the superior circular insular sulcus). The total number of bipolar iEEG derivations for the four ROIs was 44, 59, 70 and 75 for the vmPFC, lOFC, dlPFC and aINS, respectively (Fig. 1A [↗](#)). As channels with artifacts or epileptic activities were removed here, the number of recordings differs from a previous study (Gueguen et al., 2021 [↗](#)).

Behavioral task and set-up

Participants were asked to participate in a probabilistic instrumental learning task adapted from previous studies (Pessiglione et al., 2006 [↗](#); Palminteri et al., 2012 [↗](#)). Participants received written instructions that the goal of the task was to maximize their financial payoff by considering reward-seeking and punishment avoidance as equally important. Instructions were reformulated orally if necessary. Participants started with a short session, with only two pairs of cues presented on 16 trials, followed by 2–3 short sessions of 5 min. At the end of this short training, all participants were familiar with the timing of events, with the response buttons and all reached a threshold of at least 70% of correct choices during both reward and punishment conditions. Participants then performed three to six sessions on a single testing occurrence, with short breaks between sessions. Each session was an independent task, with four new pairs of cues to be learned. Cues were abstract visual stimuli taken from the Agathodaimon alphabet. The two cues of a pair were always presented together on the left and right of a central fixation cross and their

relative position was counterbalanced across trials. On each trial, one pair was randomly presented. Each pair of cues was presented 24 times for a total of 96 trials per session. The four pairs of cues were divided into two conditions. A rewarding condition where the two pairs could either lead the participants to win one euro or nothing (+1€ vs. 0€) and a symmetric punishment condition where the participants could either lose one euro or nothing (−1€ vs. 0€). Rewarding and punishing pairs of cues were presented in an intermingled random manner and participants had to learn the four pairs at once. Within each pair, the two cues were associated with the two possible outcomes with reciprocal probabilities (0.75/0.25 and 0.25/0.75). To choose between the left or right cues, participants used their left or right index to press the corresponding button on a joystick (Logitech Dual Action). Since the position on the screen was counterbalanced, response (left versus right) and value (good vs. bad cue) were orthogonal. The chosen cue was colored in red for 250 ms and then the outcome was displayed on the screen after 1000 ms. To win money, participants had to learn by trial and error which cue-outcome association was the most rewarding in the rewarding condition and the least penalizing in the punishment condition. Visual stimuli were delivered on a 19-inch TFT monitor with a refresh rate of 60 Hz, controlled by a PC with Presentation 16.5 (Neurobehavioral Systems, Albany, CA).

Computational model of learning

To model choice behavior and estimate prediction error signals, we used a standard Q-learning model (Watkins and Dayan, 1992 [\[1\]](#)) from reinforcement learning theory (Sutton and Barto, 2018 [\[2\]](#)). For a pair of cues A and B , the model estimates the expected value of choosing A (Qa) or B (Qb), given previous choices and received outcomes. Q-values were initiated to 0, corresponding to the average of all possible outcome values. After each trial t , the expected value of choosing a stimulus (e.g. A) was updated according to the following update rule:

$$Qa_{t+1} = Qa_t + \alpha \delta_t \quad (1)$$

with α the learning rate weighting the importance given to new experiences and δ , the outcome prediction error (PE) signals at a trial t defined as the difference between the obtained and expected outcomes:

$$\delta_t = R_t - Qa_t \quad (2)$$

with R_t the reinforcement value among −1€, 0€ and 1€. The probability of choosing a cue was then estimated by transforming the expected values associated with each cue using a softmax rule with a Gibbs distribution. An additional Θ parameter was added in the softmax function to the expected value of the chosen option on the previous trial of the same cue to account for the tendency to repeat the choice made on the previous trial. For example, if a participant chose option A on trial t , the probability of choosing A at trial $t+1$ was obtained using :

$$Pa_{t+1} = \frac{e^{(Qa_t + \Theta)/\beta}}{e^{(Qa_t + \Theta)/\beta} + e^{Qb_t/\beta}} \quad (3)$$

with β the choice temperature for controlling the ratio between exploration and exploitation. The three free parameters α , β and Θ were fitted per participant and optimized by minimizing the negative log-likelihood of choice using the MATLAB *fmincon* function, initialized at multiple starting points of the parameter space (Palmeri et al., 2015 [\[3\]](#)). Estimates of the free parameters, the goodness of fit and the comparison between modeled and observed data can be seen in **Table 1** [\[4\]](#) and **Figure 1** [\[5\]](#) in (Gueguen et al., 2021 [\[6\]](#)).

	T-value	P-value	P-value (FDR corrected)	dof	CI95%
$\Pi_{\text{PPE-RPE}}$	2,859	0.007**	0.015*	34	[6.5e-05, 3.9e-04]
Π_{Mixed}	2,841	0.008**	0.015*	33	[5.4e-05, 3.3e-04]
$\Pi_{\text{PPE-PPE}}$	1,25	0.2667	0.3556	5	[-7.1e-05, 2.1e-04]
$\Pi_{\text{RPE-RPE}}$	0,733	0.4912	0.4912	6	[-3.1e-05, 5.8e-05]

Table 1.
Results of the one sample t-test performed against 0

iEEG data analysis

Estimate of single-trial gamma-band activity

Here, we focused solely on broadband gamma for three main reasons. First, it has been shown that the gamma band activity correlates with both spiking activity and the BOLD fMRI signals (Mukamel et al., 2004 [↗](#); Niessing et al., 2005 [↗](#); Lachaux et al., 2007 [↗](#); Nir et al., 2007 [↗](#)), and it is commonly used in MEG and iEEG studies to map task-related brain regions (Brovelli et al., 2005 [↗](#); Crone et al., 2006 [↗](#); Vidal et al., 2006 [↗](#); Ball et al., 2008 [↗](#); Jerbi et al., 2009 [↗](#); Lachaux et al., 2012 [↗](#); Cheyne and Ferrari, 2013 [↗](#)). Therefore, focusing on the gamma band facilitates linking our results with the fMRI and spiking literatures on probabilistic learning. Second, single-trial and time-resolved high-gamma activity can be exploited for the analysis of cortico-cortical interactions in humans using MEG and iEEG techniques (Brovelli et al., 2015 [↗](#), 2017 [↗](#); Combrisson et al., 2022a [↗](#)). Finally, while previous analyses of the current dataset (Gueguen et al., 2021 [↗](#)) reported an encoding of PE signals at different frequency bands, the power in lower frequency bands were shown to carry redundant information compared to the gamma band power. In the current study, we thus estimated the power in the gamma band using a multitaper time-frequency transform based on Slepian tapers (Percival et al., 1993 [↗](#); Mitra and Pesaran, 1999 [↗](#)). To extract gamma-band activity from 50 to 100Hz, the iEEG time series were multiplied by 9 orthogonal tapers (15 cycles for a duration of 200ms and with a time-bandwidth for frequency smoothing of 10Hz), centered at 75 Hz and Fourier-transformed. To limit false negatives proportions due to multiple testings, we down-sampled the gamma power to 256 Hz. Finally, we smoothed the gamma power using a 10 points Savitzky-Golay filter. We used MNE-Python (Gramfort et al., 2013 [↗](#)) to inspect the time series, reject contacts contaminated with pathological activity, and estimate the power spectrum density (*mne.time_frequency.psd_multitaper* [↗](#)) and the gamma power (*mne.time_frequency.tfr_multitaper* [↗](#)).

Local correlates of PE signals

To quantify the local encoding of prediction error (PE) signals in the four ROIs, we used information-theoretic metrics. To this end, we computed the time-resolved mutual information (MI) between the single-trial gamma-band responses and the outcome-related PE signals. As a reminder, mutual information is defined as:

$$I(X; Y) = H(X) - H(X|Y) \quad (4)$$

In this equation, the variables X and Y represent the across-trials gamma-band power and the PE variables, respectively. $H(X)$ is the entropy of X, and $H(X|Y)$ is the conditional entropy of X given Y. In the current study, we used a semi-parametric binning-free technique to calculate MI, called Gaussian-Copula Mutual Information (GCMI) (Ince et al., 2017 [↗](#)). The GCMI is a robust rank-based approach that allows the detection of any type of monotonic relation between variables and it has been successfully applied to brain signals analysis (Colenbier et al., 2020 [↗](#); Michelmann et al., 2021 [↗](#); Ten Oever et al., 2021 [↗](#)). Mathematically, the GCMI is a lower-bound estimation of the true MI and it does not depend on the marginal distributions of the variables, but only on the copula function that encapsulates their dependence. The rank-based copula-normalization preserves the relationship between variables as long as this relation is strictly increasing or decreasing. As a consequence, the GCMI can only detect monotonic relationships. Nevertheless, the GCMI is of practical importance for brain signal analysis for several reasons. It allows to estimate the MI on a limited number of samples and it contains a parametric bias correction to compensate for the bias due to the estimation on smaller datasets. It allows to compute the MI on uni- and multivariate variables that can either be continuous or discrete (see **Table I** [↗](#) in (Ince et al., 2017 [↗](#))). Finally, it is computationally efficient, which is a desired property when dealing with a

large number of iEEG contacts recording at a high sampling rate. Here, the GCMI was computed across trials and it was used to estimate the instantaneous amount of information shared between the gamma power of iEEG contacts and RPE ($MI_{RPE} = I(\gamma; RPE)$) and PPE signals ($M_{PPE} = I(\gamma; PPE)$).

Network-level interactions and PE signals

The goal of network-level analyses was to characterize the nature of cortico-cortical interactions encoding reward and punishment PE signals. In particular, we aimed to quantify: 1) the nature of the interdependence between pairs of brain ROIs in the encoding of PE signals; 2) the information flow between ROIs encoding PE signals. These two questions were addressed using Interaction Information and Transfer Entropy analyses, respectively.

Interaction Information analysis

In classical information theory, interaction information (II) provides a generalization of mutual information for more than two variables (McGill, 1954; Ince et al., 2017). For the three-variables case, the II can be defined as the difference between the total, or joint, mutual information between ROIs (R_1 and R_2) and the third behavioral variable (S), minus the two individual mutual information between each ROI and the behavioral variable. For a three variables multivariate system composed of two sources R_1 , R_2 and a target S , the II is defined as :

$$\begin{aligned} II(R_1; R_2; S) &= I(S; R_1 | R_2) - I(R_1; S) \\ &= I(R_1, R_2; S) - I(R_1; S) - I(R_2; S) \end{aligned} \quad (5)$$

Unlike mutual information, the interaction information can be either positive or negative. A negative value of interaction information indicates a net redundant effect between variables, whereas positive values indicate a net synergistic effect (Williams and Beer, 2010). Here, we used the II to investigate the amount of information and the nature of the interactions between the gamma power of pairs of contacts (γ_1, γ_2) about the RPE ($II_{RPE} = II(\gamma_1, \gamma_2; RPE)$) and PPE signals ($II_{PPE} = II(\gamma_1, \gamma_2; PPE)$). The II was computed by estimating the MI quantities of equation (5) using the GCMI between contacts within the same brain region or across different regions.

Transfer Entropy analysis

To quantify the degree of communication between neural signals, the most successful model-free methods rely on the Wiener-Granger principle (Wiener, 1956; Granger, 1969). This principle identifies information flow between time series when future values of a given signal can be predicted from the past values of another, above and beyond what can be achieved from its autocorrelation. One of the most general information theoretic measures based on the Wiener-Granger principle is Transfer Entropy (TE) (Schreiber, 2000). The TE can be formulated in terms of conditional mutual information (Schreiber, 2000; Kaiser and Schreiber, 2002):

$$TE(X \rightarrow Y) = I(X_{past}; Y_t | Y_{past}) \quad (6)$$

Here, we computed the TE on the gamma activity time courses of pairs of iEEG contacts. We used the GCMI to estimate conditional mutual information. For an interval $[d_1, d_2]$ of n_{delays} , the final TE estimation was defined as the mean over the TE estimated at each delay :

$$TE(X \rightarrow Y)_{[d_1, d_2]} = \frac{1}{n_{delays}} \cdot \sum_{d=d_1}^{d_2} I(X_d; Y_t | Y_d) \quad (7)$$

Statistical analysis

We used a group-level approach based on non-parametric permutations, encompassing non-negative measures of information (Combrisson et al., 2022a [↗](#)). The same framework was used at the local level (i.e. the information carried by single contact) or at the network level (i.e. the information carried by pairs of contacts for the II and TE). To take into account the inherent variability existing at the local and network levels, we used a random-effect model. To generate the distribution of permutations at the local level, we shuffled the PE variable across trials 1000 times and computed the MI between the gamma power and the shuffled version of the PE. The shuffling led to a distribution of MI reachable by chance, for each contact and at each time point (Combrisson and Jerbi, 2015 [↗](#)). To form the group-level effect, we computed a one-sample t-test against the permutation mean across the MI computed on individual contacts taken from the same brain region, at each time point. The same procedure was used on the permutation distribution to form the group-level effect reachable by chance. We used cluster-based statistics for correcting for multiple comparisons (Maris and Oostenveld, 2007 [↗](#)). The cluster-forming threshold was defined as the 95th percentile of the distribution of t-values obtained from the permutations. We used this threshold to form the temporal clusters within each brain region. We obtained cluster masses on both the true t-values and the t-values computed on the permutations. To correct for multiple comparisons, we built a distribution made of the largest 1000 cluster masses estimated on the permuted data. The final corrected p-values were inferred as the proportion of permutations exceeding the t-values. To generate the distributions of II and TE reachable by chance, we respectively shuffled the PE variable across trials for the II and the gamma power across trials of the source for the TE (Vicente et al., 2011 [↗](#)). The rest of the significance testing procedure at the network level is similar to the local level, except that it is not applied within brain regions but within pairs of brain regions.

Software

Information-theoretic metrics and group-level statistics, are implemented in a homemade Python software called *Frites* (Combrisson et al., 2022b [↗](#)). The interaction information can be computed using the `frites.conn.conn_ii` [↗](#) function and the transfer entropy using the `frites.conn.conn_te` [↗](#) function.

Acknowledgements

We would like to express our gratitude to Benjamin Morillon, Manuel R. Mercier, and Stefano Palminteri for their valuable comments on an earlier draft of this manuscript and for their feedback on our responses to reviewer comments.

Code and data availability

The Python scripts and notebooks to reproduce the results presented here are hosted on Github (https://github.com/brainets/papercode/tree/main/combrisson_pblt_2024 [↗](#)). The preprocessed data used here can be downloaded using the following link : <https://amubox.univ-amu.fr/s/Mwdy3NiE8FdrW6W> [↗](#)

Supplementary figures

Figure S1.

Single subject anatomical repartition

(A) Number of unique subject per brain region and per pair of brain regions (B) Number of bipolar derivations per subject and per brain region

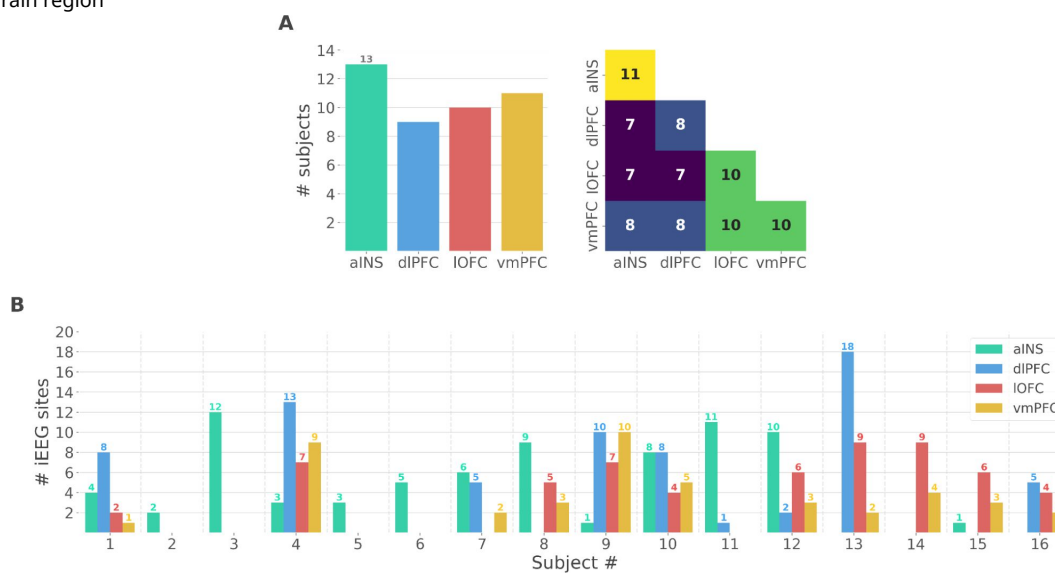
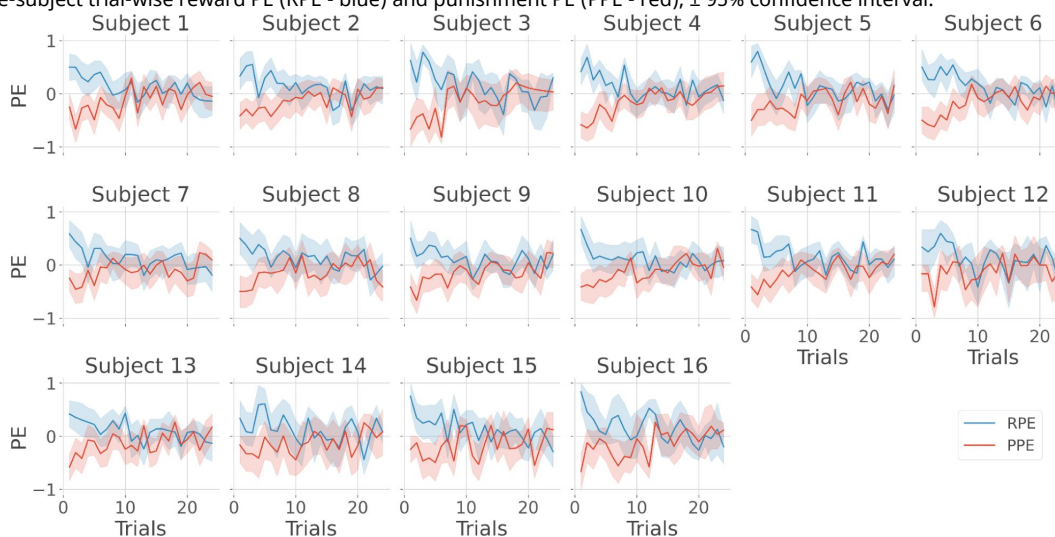


Figure S2.

Single-subject estimation of predictions errors

Single-subject trial-wise reward PE (RPE - blue) and punishment PE (PPE - red), $\pm$ 95% confidence interval.



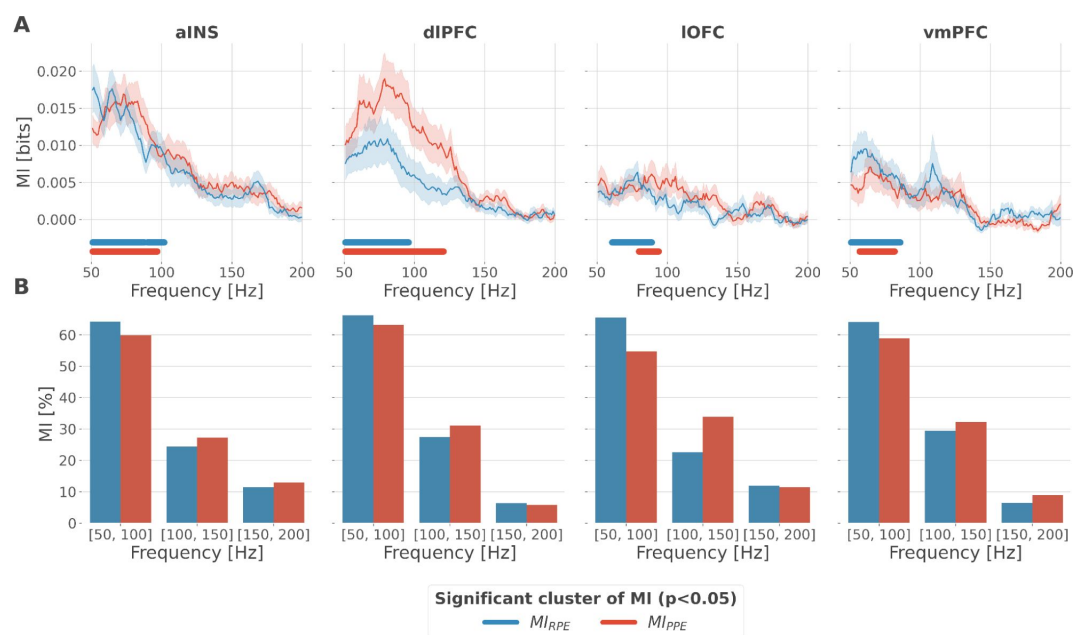


Figure S3.

Local encoding of prediction error signals within the gamma band

We searched for the most informative sub-band about the R/PPE within the broad gamma range. To this end, we estimated the PSD between [50, 200]Hz during the first second following the outcome apparition. We used multitapers with a 20 Hz bandwidth for the multi-taper windowing function. We then estimated the amount of information carried by individual gamma frequencies about the R/PPE in the four brain regions. We found significant clusters of MI over the four brain regions with both RPE and PPE signals mostly within the [50, 100]Hz range (**Fig. S1A**). We then estimated the density of information within the [50, 100]Hz, [100, 150]Hz and [150, 200]Hz bins (**Fig. S1B**). Approximately 60% of the total information was concentrated within the [50, 100]Hz range. (A) Distribution of information in the aINS, dIPFC, IOFC and vmPFC about the R/PPE in the frequency domain. The solid line and the shaded area respectively represent the mean and SEM of the across-contacts MI. Horizontal thick lines represent significant clusters of information ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs), (B) Density of information in the [50, 100]Hz, [100, 150]Hz and [150, 200]Hz bins.

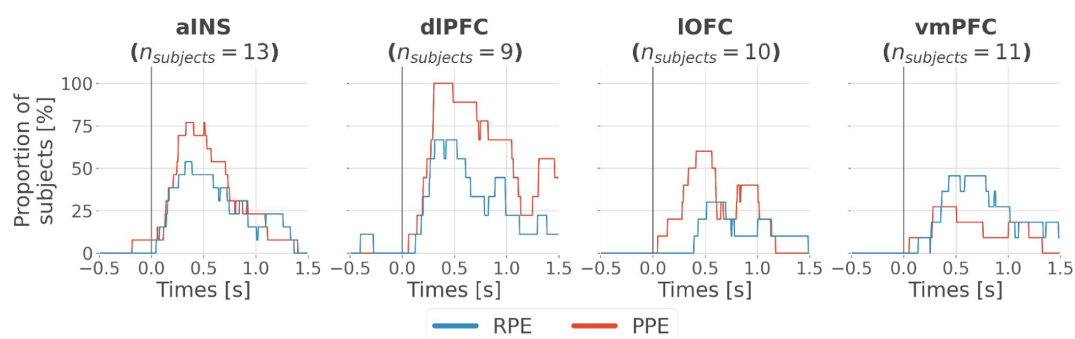


Figure S4.

Inter-subjects reproducibility of local encoding of PE signals

To assess that the local encoding of RPE and PPE signals were not driven by a minority of subjects, we computed the mutual information between the gamma activity of each bipolar derivation and the R/PPE and performed statistical inferences at the single-subject level (correction for multiple comparisons across time points and bipolar derivations). We then estimated the proportion of unique subjects per brain region with at least one significant bipolar derivation (**Fig. S4**). We found a minimum proportion of approximately 30% reproducibility achieved in the IOFC and vmPFC for the encoding of the RPE and PPE, respectively. Conversely, between 75-100% of the subjects having bipolar derivations in the aINS and dlPFC presented a significant encoding of the PPE. Time-courses of proportion of subjects having at least one bipolar derivation with a significant encoding ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs) of RPE (blue) or PPE (red). Data is aligned to the outcome presentation (vertical line at 0 seconds).

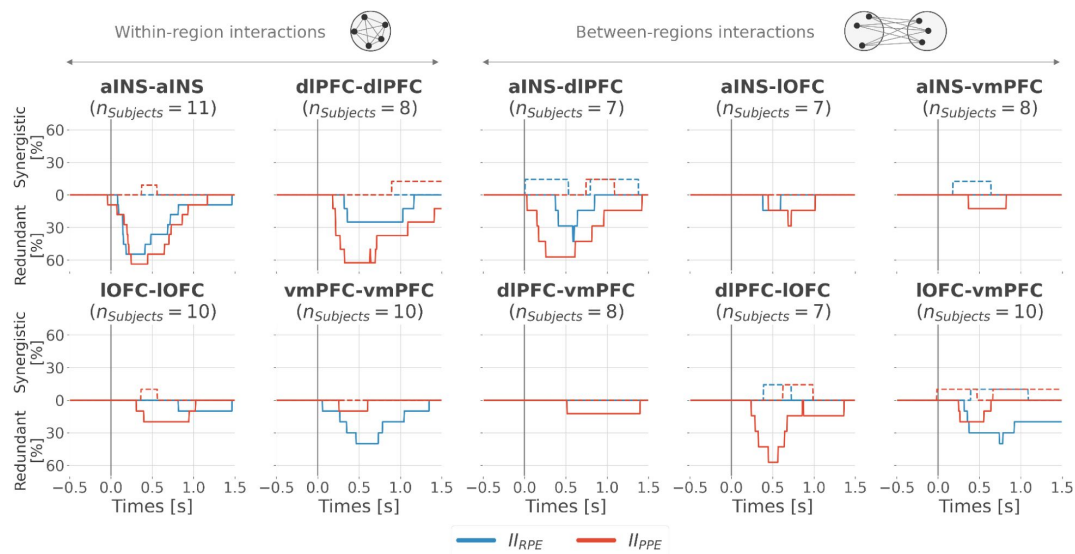


Figure S5.

Inter-subjects reproducibility of redundant interactions about PE signals

Similarly, we quantified the reproducibility of the type of interactions the four brain regions had about the RPE and PPE. We computed the interaction information between all pairs of bipolar derivations per subject and performed the statistical inferences at the single-subject level (correction for multiple comparisons across time points and pairs bipolar derivations). We finally estimated the proportion of unique subjects per pair of brain regions having at least one significant pair of bipolar derivations with either redundant or synergistic interactions with either the RPE and PPE (Fig. S5). For the within-regions interactions, approximately 60% of the subjects had redundant interactions about R/PPE in the aINS and about the PPE in the dIPFC and 40% about the RPE in the vmPFC. For the across-regions interactions, 60% of the subjects had redundant interactions between the aINS-dIPFC and dIPFC-IOFC about the PPE, and 30% had redundant interactions between IOFC-vmPFC about the RPE. Time-courses of proportion of subjects having at least one pair of bipolar derivation with significant interaction information ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs) about the RPE (blue) or PPE (red). Data is aligned to the outcome presentation (vertical line at 0 seconds). Proportion of subjects with redundant (solid) and synergistic (dashed) interactions are respectively going downward and upward.

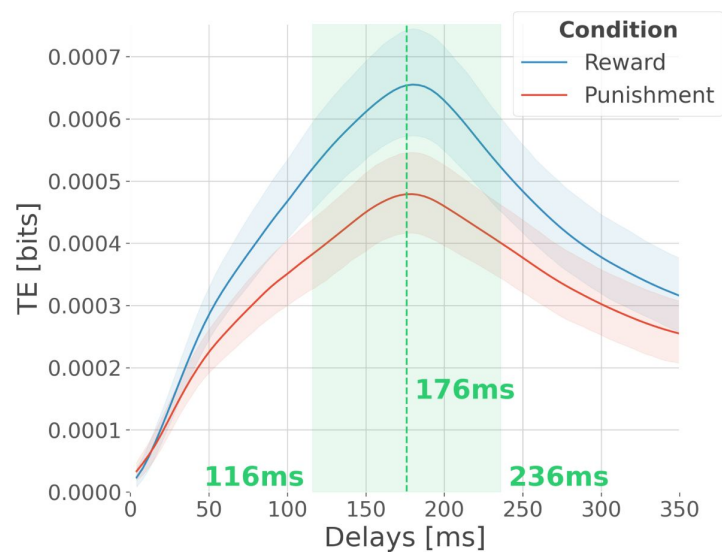


Figure S6.

Optimal delay interval for maximising information transfer

We searched for the optimal delay interval to maximise the information transfer between regions. An important parameter to consider when estimating TE is the number of time points in the past of the target to use for conditioning. An information flow from source X to target Y exists because the inclusion of the past of X reduces the uncertainty about the future of Y, given its own past. We investigated the influence of the delay on the TE. To this end, we estimated the TE across all pairs of contacts per participant, for rewarding and punishing trials, at every possible delay up to 350ms (**Fig. S6**). We found a maximum information flow for delays up to 176 ms. Therefore, for the main text analyses, we used a range centred around 176 ± 60 ms ([116, 236]ms). Modulation of transfer entropy (TE in bits) as a function of the delay between source and target areas. In blue, the TE computed across rewarding trials and in red, the TE computed across punishing trials. Shaded areas surrounding the time courses represent the 95% confidence interval estimated using a bootstrapping strategy.

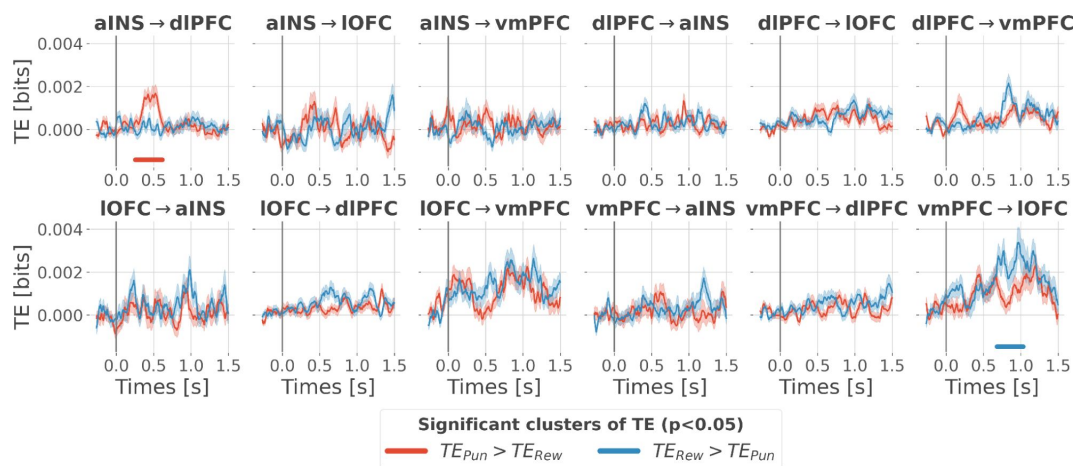


Figure S7.

Differential cortico-cortical directional interactions

We estimated the transfer entropy (TE) on the gamma power during the rewarding (TE_{Rew}) and punishment conditions (TE_{Pun}). As in the main text, we computed the TE for all pairs of brain regions within delays between 116 and 236 msec and detected temporal clusters where the TE significantly differed between conditions (TE_{Rew} > TE_{Pun} or TE_{Pun} > TE_{Rew}). Only two pairs of brain regions displayed statistically significant modulations in TE (**Fig. S7**). The TE from the aINS to the dIPFC (TE_{aINS → dIPFC}) during the punishment condition and the TE from the vmPFC to the IOFC (TE_{vmPFC → IOFC}) during the rewarding condition. No other brain interactions were found significant. Time courses of transfer entropy (TE, in bits) estimated during the rewarding condition (TE_{Rew} in blue) and punishing condition (TE_{Pun} in red). Significant differences (p < 0.05, cluster-based correction, non-parametric randomization across epochs) of TE between conditions are displayed with horizontal bold lines (blue for TE_{Rew} > TE_{Pun} and red for TE_{Pun} > TE_{Rew}). Shaded areas represent the SEM. The vertical grey line at 0 seconds represents the outcome presentation.

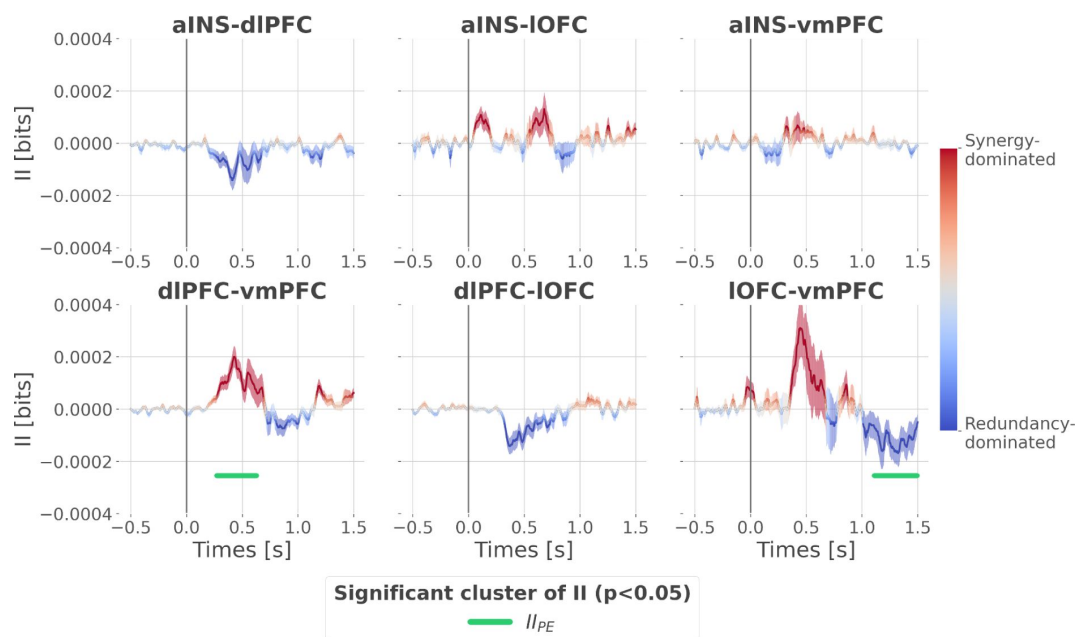


Figure S8.

Cortico-cortical interactions about the full PE signals

Dynamic interaction information (II in bits) between-regions about the full prediction error (II_{PE}). Hot and cold colours indicate synergy- and redundancy-dominated interactions about the full PE. Significant clusters of II_{PE} are displayed with a horizontal bold green line ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs). Shaded areas represent the SEM. The vertical grey line at 0 seconds represents the outcome presentation.

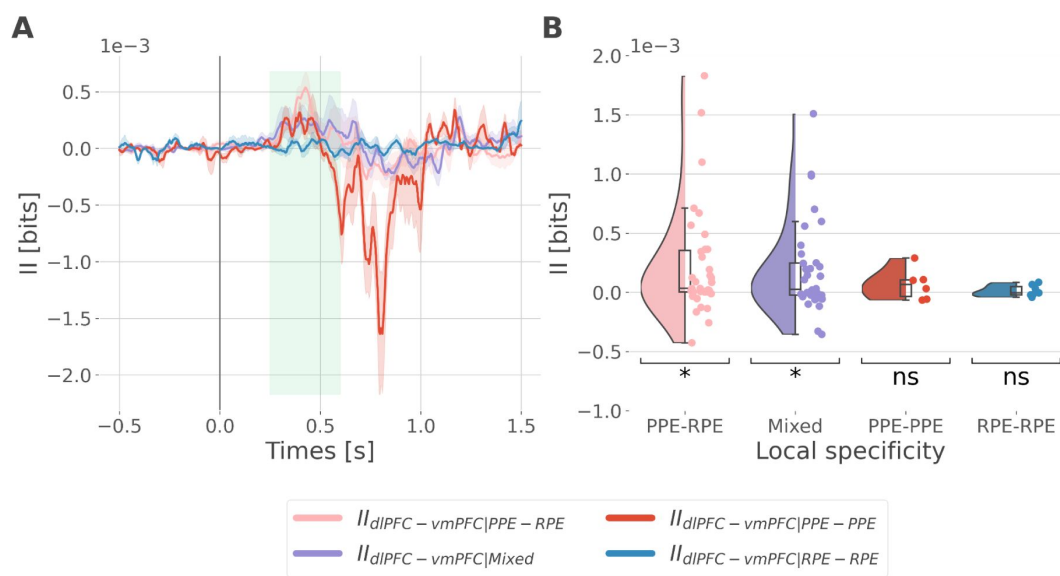


Figure S9.

Interaction information binned according to the local specificity

We binned the II about the full PE (i.e. by concatenating the RPE and PPE) according to the local specificity of the bipolar derivations in the dIPFC and vmPFC i.e. contacts with gamma activity modulated according to the RPE only, to the PPE only or to both RPE and PPE (**Fig. 2B**). As a result, we binned the II into four categories: the IIRPE-RPE and IIPPE-PPE respectively reflecting the II estimated between recordings specific to the RPE and PPE, the IIPPE-RPE between recordings PPE and RPE specific and the IIMixed for the remaining possibilities (i.e. RPE-Both, PPE-Both and Both-Both) (**Fig. S9A**). We reported a significant cluster II about the full PE approximately between [250; 600]ms (**Fig. 5**). Therefore, we estimated the mean II across time points between [250; 600]ms within each of the four categories, for each pair of recordings (**Fig. S9B**) and computed a one-sample t-test against 0 (**Table 1**). Only the $II_{PPE-RPE}$ (pink) and II_{Mixed} (purple) mean showed a significant difference from 0. However, the number of pairs of recordings for the $II_{RPE-RPE}$ (red, 7 pairs) and $II_{PPE-PPE}$ (blue, 6 pairs) were probably too small to find a significant difference. (A) Dynamic interaction information (II in bits) between the dIPFC and vmPFC ($II_{dIPFC - vmPFC}$) binned according to the local specificity toward the RPE and PPE. Shaded areas represent the SEM. The vertical grey line at 0 seconds represents the outcome presentation. (B) Mean II between time points from 250 to 600 ms after outcome presentation per category of local specificity. Each individual point represents one pair of recordings from the dIPFC and vmPFC.

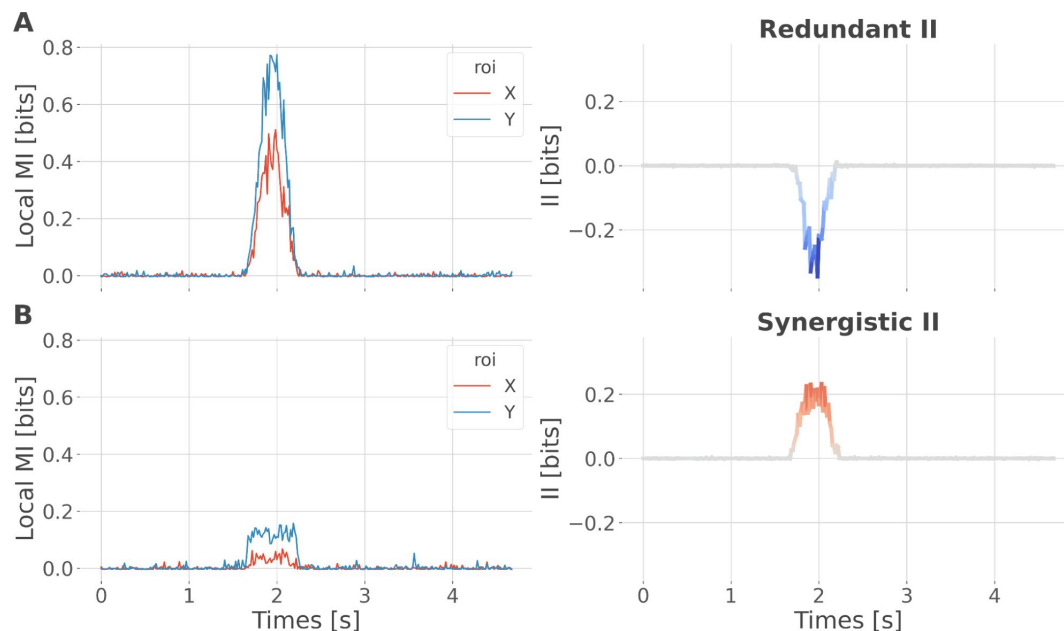


Figure S10.

Simulation study testing for synergistic effects

We performed a simulation to demonstrate that synergistic interactions can emerge between two regions with the same specificity. For example, consider one region that locally encodes early trials of reward prediction error (RPE) and a second region that encodes late trials of RPE. Combining the two using the interaction information (II) measure would lead to synergistic interactions, as each region carries information that is not carried by the other.

To simulate this scenario, we initialized data for two brain regions, X and Y, and a 200-trial prediction error vector, all using random noise sampled from a uniform distribution. To simulate redundant interactions (**Fig. S10A**), both X and Y received a copy of the prediction error (one-to-all). To simulate synergy (**Fig. S10B**), X and Y received early and late prediction error trials, respectively (all-to-one).

In both scenarios, local mutual information (MI) encoding the PE increased for regions X and Y around 1.5 seconds. However, in the first case, it led to negative II (redundancy), while in the second case, it led to positive II (synergy). This toy example illustrates that local specificity is not the only factor determining the type of interactions between regions.

Within-area local encoding of PE using the mutual information (MI, in bits) for regions X and Y and between-area interaction information (II, in bits) leading to (A) redundant interactions and (B) synergistic interactions about the PE.

References

- Auzias G, Coulon O, Brovelli A (2016) **MarsAtlas : A cortical parcellation atlas for functional mapping: MarsAtlas** *Hum Brain Mapp* **37**:1573–1592
- Averbeck B, O'Doherty JP (2022) **Reinforcement-learning in fronto-striatal circuits** *Neuropsychopharmacology* **47**:147–162
- Averbeck BB, Latham PE, Pouget A (2006) **Neural correlations, population coding and computation** *Nat Rev Neurosci* **7**:358–366
- Averbeck BB, Murray EA (2020) **Hypothalamic Interactions with Large-Scale Neural Circuits Underlying Reinforcement Learning and Motivated Behavior** *Trends Neurosci* **43**:681–694
- Ball T, Demandt E, Mutschler I, Neitzel E, Mehring C, Vogt K, Aertsen A, Schulze-Bonhage A (2008) **Movement related activity in the high gamma range of the human EEG** *Neuroimage* **41**:302–310
- Balleine BW (2019) **The Meaning of Behavior: Discriminating Reflex and Volition in the Brain** *Neuron* **104**:47–62
- Balleine BW, Dickinson A (1998) **Goal-directed instrumental action: contingency and incentive learning and their cortical substrates** *Neuropharmacology* **37**:407–419
- Balleine BW, O'doherty JP (2010) **Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action** *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol* **35**:48–69
- Barlow H (2001) **Redundancy reduction revisited** *Netw Comput Neural Syst* **12**
- Bartra O, McGuire JT, Kable JW (2013) **The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value** *NeuroImage* **76**:412–427
- Bassett DS, Mattar MG (2017) **A Network Neuroscience of Human Learning: Potential to Inform Quantitative Theories of Brain and Behavior** *Trends Cogn Sci* **21**:250–264
- Bastin J *et al.* (2016) **Direct Recordings from Human Anterior Insula Reveal its Leading Role within the Error-Monitoring Network** *Cereb Cortex:bhv* **352**
- Battaglia D, Brovelli A (2020) **Functional connectivity and neuronal dynamics: insights from computational methods** *Cogn Neurosci Sixth Ed*
- Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, Salzman CD (2020) **The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex** *Cell* **183**:954–967
- Bódi N, Kéri S, Nagy H, Moustafa A, Myers CE, Daw N, Dibó G, Takats A, Bereczki D, Gluck MA (2009) **Reward-learning and the novelty-seeking personality: a between-and within-subjects study of the effects of dopamine agonists on young Parkinson's patients** *Brain J Neurol* **132**:2385–2395

Bouton ME (2007) **Learning and behavior: A contemporary synthesis**

Bressler SL, Menon V (2010) **Large-scale brain networks in cognition: emerging methods and principles** *Trends Cogn Sci* **14**:277–290

Brovelli A, Badier J-M, Bonini F, Bartolomei F, Coulon O, Auzias G (2017) **Dynamic Reconfiguration of Visuomotor-Related Functional Connectivity Networks** *J Neurosci* **37**:839–853

Brovelli A, Lachaux J-P, Kahane P, Boussaoud D (2005) **High gamma frequency oscillatory activity dissociates attention from intention in the human premotor cortex** *NeuroImage* **28**:154–164

Brovelli XA, Chicharro D, Badier J-M, Wang H, Jirsa V (2015) **Characterization of Cortical Networks and Corticocortical Functional Connectivity Mediating Arbitrary Visuomotor Mapping** *J Neurosci* **35**:12643–12658

Buehlmann A, Deco G, Friston KJ (2010) **Optimal Information Transfer in the Cortex through Synchronization** *PLoS Comput Biol* **6**

Buzsáki G, Draguhn A (2004) **Neuronal oscillations in cortical networks** *science* **304**:1926–1929

Carmichael ST, Price J (1996) **Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys** *J Comp Neurol* **371**:179–207

Cheyne D, Ferrari P (2013) **MEG studies of motor cortex gamma oscillations: evidence for a gamma “fingerprint” in the brain?** *Front Hum Neurosci* **7**

Cloutman LL, Binney RJ, Drakesmith M, Parker GJM, Lambon Ralph MA (2012) **The variation of function across the human insula mirrors its patterns of structural connectivity: Evidence from in vivo probabilistic tractography** *NeuroImage* **59**:3514–3521

Colenbier N, Van de Steen F, Uddin LQ, Poldrack RA, Calhoun VD, Marinazzo D (2020) **Disambiguating the role of blood flow and global signal with partial information decomposition** *Neuroimage* **213**

Combrisson E, Allegra M, Basanisi R, Ince RAA, Giordano BL, Bastin J, Brovelli A (2022) **Group-level inference of information-based measures for the analyses of cognitive brain networks from neurophysiological data** *NeuroImage* **258**

Combrisson E, Basanisi R, Cordeiro VL, Ince RAA, Brovelli A (2022) **Frites: A Python package for functional connectivity analysis and group-level statistics of neurophysiological data** *Open Source Softw* **7**

Combrisson E, Jerbi K (2015) **Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy** *J Neurosci Methods* **250**:126–136

Combrisson E, Perrone-Bertolotti M, Soto JL, Alamian G, Kahane P, Lachaux J-P, Guillot A, Jerbi K (2017) **From intentions to actions: Neural oscillations encode motor processes through phase, amplitude and phase-amplitude coupling** *NeuroImage* **147**:473–487

- Crone NE, Sinai A, Korzeniewska A (2006) **High-frequency gamma oscillations and human brain mapping with electrocorticography** *Progress in Brain Research* :275–295
- D’Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008) **BOLD responses reflecting dopaminergic signals in the human ventral tegmental area** *Science* **319**:1264–1267
- Deman P, Bhattacharjee M, Tadel F, Job A-S, Rivière D, Cointepas Y, Kahane P, David O (2018) **IntrAnat electrodes: a free database and visualization software for intracranial electroencephalographic data processed for case and group studies** *Front Neuroinformatics* **12**
- Destrieux C, Fischl B, Dale A, Halgren E (2010) **Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature** *Neuroimage* **53**:1–15
- Dickinson A, Balleine B (1994) **Motivational control of goal-directed action** *Anim Learn Behav* **22**:1–18
- Diekhof EK, Kaps L, Falkai P, Gruber O (2012) **The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude—An activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing** *Neuropsychologia* **50**:1252–1266
- Dolan RJ, Dayan P (2013) **Goals and habits in the brain** *Neuron* **80**:312–325
- Engel AK, Fries P, Singer W (2001) **Dynamic predictions: oscillations and synchrony in top-down processing** *Nat Rev Neurosci* **2**:704–716
- Fedorenko E, Thompson-Schill SL (2014) **Reworking the language network** *Trends Cogn Sci* **18**:120–126
- Fouragnan E, Retzler C, Philiastides MG (2018) **Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis** *Hum Brain Mapp* **39**:2887–2906
- Frank MJ, Seeberger LC, O’reilly RC (2004) **By carrot or by stick: cognitive reinforcement learning in parkinsonism** *Science* **306**:1940–1943
- Fries P (2015) **Rhythms for Cognition: Communication through Coherence** *Neuron* **88**:220–235
- Fusi S, Miller EK, Rigotti M (2016) **Why neurons mix: high dimensionality for higher cognition** *Curr Opin Neurobiol* **37**:66–74
- Garrison J, Erdeniz B, Done J (2013) **Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies** *Neurosci Biobehav Rev* **37**:1297–1310
- Gelens F, Komatsu M, Uran C, Jensen MA, Miller KJ, Ince RA, Vinck M, Canales-Johnson A (2023) **Distributed representations of prediction error signals across the cortical hierarchy are synergistic** *BioRxiv Prepr Serv Biol* :2023–1
- Ghaziri J, Tucholka A, Girard G, Houde J-C, Boucher O, Gilbert G, Descoteaux M, Lippé S, Rainville P, Nguyen DK (2017) **The Corticocortical Structural Connectivity of the Human Insula** *Cereb Cortex* **27**:1216–1228

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L (2013) **MEG and EEG data analysis with MNE-Python** *Front Neurosci* **7**

Granger CW (1969) **Investigating causal relations by econometric models and cross-spectral methods** *Econom J Econom Soc* :424–438

Gueguen MCM, Lopez-Persem A, Billeke P, Lachaux J-P, Rheims S, Kahane P, Minotti L, David O, Pessiglione M, Bastin J (2021) **Anatomical dissociation of intracerebral signals for reward and punishment prediction errors in humans** *Nat Commun* **12**

Gutnisky DA, Dragoi V (2008) **Adaptive coding of visual information in neural populations** *Nature* **452**:220–224

Hsu C-C Heather, Rolls ET, Huang C-C, Chong ST, Zac Lo C-Y, Feng J, Lin C-P (2020) **Connections of the Human Orbitofrontal Cortex and Inferior Frontal Gyrus** *Cereb Cortex* **30**:5830–5843

Helfrich RF, Knight RT (2019) **Cognitive neurophysiology of the prefrontal cortex** *Handbook of Clinical Neurology* :35–59

Hirokawa J, Vaughan A, Masset P, Ott T, Kepecs A (2019) **Frontal cortex neuron types categorically encode single decision variables** *Nature* **576**:446–451

Hunt LT, Hayden BY (2017) **A distributed, hierarchical and recurrent framework for reward-based choice** *Nat Rev Neurosci* **18**:172–182

Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, Behrens TE (2012) **Mechanisms underlying cortical activity during value-guided choice** *Nat Neurosci* **15**:470–476

Ince RAA, Giordano BL, Kayser C, Rousselet GA, Gross J, Schyns PG (2017) **A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula: Gaussian Copula Mutual Information** *Hum Brain Mapp* **38**:1541–1573

Jerbi K *et al.* (2009) **Task-related gamma-band dynamics from an intracerebral perspective: Review and implications for surface EEG and MEG** *Hum Brain Mapp* **30**:1758–1771

Jocham G, Hunt LT, Near J, Behrens TE (2012) **A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex** *Nat Neurosci* **15**:960–961

Kafashan M, Jaffe AW, Chettih SN, Nogueira R, Arandia-Romero I, Harvey CD, Moreno-Bote R, Drugowitsch J (2021) **Scaling of sensory information in large neural populations shows signatures of information-limiting correlations** *Nat Commun* **12**

Kaiser A, Schreiber T (2002) **Information transfer in continuous processes** *Phys Nonlinear Phenom* **166**:43–62

Kirst C, Timme M, Battaglia D (2016) **Dynamic information routing in complex networks** *Nat Commun* **7**

Lachaux J-P, Axmacher N, Mormann F, Halgren E, Crone NE (2012) **High-frequency neural activity and human cognition: past, present and possible future of intracranial EEG research** *Prog Neurobiol* **98**:279–301

- Lachaux J-P, Fonlupt P, Kahane P, Minotti L, Hoffmann D, Bertrand O, Baciau M (2007) **Relationship between task-related gamma oscillations and BOLD signal: new insights from combined fMRI and intracranial EEG** *Hum Brain Mapp* **28**:1368–1375
- Lachaux JP, Rudrauf D, Kahane P (2003) **Intracranial EEG and human brain mapping** *J Physiol-Paris* **97**:613–628
- Liu X, Hairston J, Schrier M, Fan J (2011) **Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies** *Neurosci Biobehav Rev* **35**:1219–1236
- Lizier JT, Bertschinger N, Jost J, Wibral M (2018) **Information decomposition of target effects from multi-source interactions: Perspectives on previous, current and future work** *Entropy* **20**
- Loued-Khenissi L, Pfeuffer A, Einhäuser W, Preusschoff K (2020) **Anterior insula reflects surprise in value-based decision-making and perception** *NeuroImage* **210**
- Luppi AI, Mediano PAM, Rosas FE, Holland N, Fryer TD, O’Brien JT, Rowe JB, Menon DK, Bor D, Stamatakis EA (2022) **A synergistic core for human brain evolution and cognition** *Nat Neurosci* **25**:771–782
- Luppi AI, Rosas FE, Mediano PAM, Menon DK, Stamatakis EA (2024) **Information decomposition and the informational architecture of the brain** *Trends Cogn Sci:S*
- Maris E, Oostenveld R (2007) **Nonparametric statistical testing of EEG- and MEG-data** *J Neurosci Methods* **164**:177–190
- Matsumoto M, Hikosaka O (2009) **Two types of dopamine neuron distinctly convey positive and negative motivational signals** *Nature* **459**:837–841
- McGill W (1954) **Multivariate information transmission** *Trans IRE Prof Group Inf Theory* **4**:93–111
- Mercier MR, Dubarry A-S, Tadel F, Avanzini P, Axmacher N, Cellier D, Del Vecchio M, Hamilton LS, Hermes D, Kahana MJ (2022) **Advances in human intracranial electroencephalography research, guidelines and good practices** *NeuroImage* **119438**
- Meyers EM, Freedman DJ, Kreiman G, Miller EK, Poggio T (2008) **Dynamic population coding of category information in inferior temporal and prefrontal cortex** *J Neurophysiol* **100**:1407–1419
- Michelmann S, Price AR, Aubrey B, Strauss CK, Doyle WK, Friedman D, Dugan PC, Devinsky O, Devore S, Flinker A (2021) **Moment-by-moment tracking of naturalistic learning and its underlying hippocampo-cortical interactions** *Nat Commun* **12**:1–15
- Miller EK, Brincat SL, Roy JE (2024) **Cognition is an emergent property** *Curr Opin Behav Sci* **57**
- Mitra PP, Pesaran B (1999) **Analysis of dynamic brain imaging data** *Biophys J* **76**:691–708
- Monosov IE, Hikosaka O (2012) **Regionally Distinct Processing of Rewards and Punishments by the Primate Ventromedial Prefrontal Cortex** *J Neurosci* **32**:10318–10330

- Morrison SE, Salzman CD (2009) **The Convergence of Information about Rewarding and Aversive Stimuli in Single Neurons** *J Neurosci* **29**:11471–11483
- Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R (2004) **Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex** *Cereb Cortex* **14**
- Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA (2005) **Hemodynamic signals correlate tightly with synchronized gamma oscillations** *science* **309**:948–951
- Nigam S, Pojoga S, Dragoi V (2019) **Synergistic Coding of Visual Information in Columnar Networks** *Neuron* **104**:402–411
- Nir Y, Fisch L, Mukamel R, Gelbard-Sagiv H, Arieli A, Fried I, Malach R (2007) **Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations** *Curr Biol* **17**:1275–1285
- Noble S, Curtiss J, Pessoa L, Scheinost D (2024) **The tip of the iceberg: A call to embrace anti-localizationism in human neuroscience research** *Imaging Neurosci* **2**:1–10
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) **Dissociable roles of ventral and dorsal striatum in instrumental conditioning** *science* **304**:452–454
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) **Abstract reward and punishment representations in the human orbitofrontal cortex** *Nat Neurosci* **4**:95–102
- Ohnuki T, Osako Y, Manabe H, Sakurai Y, Hirokawa J (2021) **Over-representation of fundamental decision variables in the prefrontal cortex underlies decision bias** *Neurosci Res* **173**:1–13
- Öngür D, Price JL (2000) **The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans** *Cereb Cortex* **10**:206–219
- Palmigiano A, Geisel T, Wolf F, Battaglia D (2017) **Flexible information routing by transient synchrony** *Nat Neurosci* **20**:1014–1022
- Palminteri S *et al.* (2012) **Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning** *Neuron* **76**:998–1009
- Palminteri S, Khamassi M, Joffily M, Coricelli G (2015) **Contextual modulation of value signals in reward and punishment learning**
- Palminteri S, Lebreton M, Worbe Y, Grabli D, Hartmann A, Pessiglione M (2009) **Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes** *Proc Natl Acad Sci* **106**:19179–19184
- Palminteri S, Pessiglione M (2017) **Opponent Brain Systems for Reward and Punishment Learning** *Decision Neuroscience* :291–303
- Panzeri S, Macke JH, Gross J, Kayser C (2015) **Neural population coding: combining insights from microscopic and mass signals** *Trends Cogn Sci* **19**:162–172
- Panzeri S, Moroni M, Safaai H, Harvey CD (2022) **The structures and functions of correlations in neural population codes** *Nat Rev Neurosci* **23**:551–567

- Parras GG, Nieto-Diego J, Carbajal GV, Valdés-Baizabal C, Escera C, Malmierca MS (2017) **Neurons along the auditory pathway exhibit a hierarchical organization of prediction error** *Nat Commun* **8**
- Parthasarathy A, Herikstad R, Bong JH, Medina FS, Libedinsky C, Yen S-C (2017) **Mixed selectivity morphs population codes in prefrontal cortex** *Nat Neurosci* **20**:1770–1779
- Percival DB, Walden AT (1993) **Spectral analysis for physical applications**
- Pessiglione M, Delgado MR (2015) **The good, the bad and the brain: neural correlates of appetitive and aversive values underlying decision making** *Curr Opin Behav Sci* **5**:78–84
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) **Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans** *Nature* **442**:1042–1045
- Petersen SE, Sporns O (2015) **Brain networks and cognitive architectures** *Neuron* **88**:207–219
- Plassmann H, O'Doherty JP, Rangel A (2010) **Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making** *J Neurosci* **30**:10799–10808
- Bartolo Ramon, Bartolo R, Saunders RC, Mitz AR, Averbeck BB (2020) **Information-Limiting Correlations in Large Neural Populations** *J Neurosci* **40**:1668–1678
- Reid AT *et al.* (2019) **Advancing functional connectivity research from association to causation** *Nat Neurosci* **22**:1751–1760
- Rescorla RA, Wagner AR, Black AH, Prokasy WF (1972) **Classical conditioning II: current research and theory**
- Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, Fusi S (2013) **The importance of mixed selectivity in complex cognitive tasks** *Nature* **497**:585–590
- Saez I, Lin J, Stolk A, Chang E, Parvizi J, Schalk G, Knight RT, Hsu M (2018) **Encoding of Multiple Reward-Related Computations in Transient and Sustained High-Frequency Activity in Human OFC** *Curr Biol* **28**:2889–2899
- Saleem AB, Diamanti EM, Fournier J, Harris KD, Carandini M (2018) **Coherent encoding of subjective spatial position in visual cortex and hippocampus** *Nature* **562**:124–127
- Salinas E, Sejnowski TJ (2001) **Correlated neuronal activity and the flow of neural information** *Nat Rev Neurosci* **2**:539–550
- Schneider M, Broggin AC, Dann B, Tzanou A, Uran C, Sheshadri S, Scherberger H, Vinck M (2021) **A mechanism for inter-areal coherence through communication based on connectivity and oscillatory power** *Neuron* **109**:4050–4067
- Schneidman E, Bialek W, Berry MJ (2003) **Synergy, Redundancy, and Independence in Population Codes** *J Neurosci* **23**:11539–11553
- Schreiber T (2000) **Measuring information transfer** *Phys Rev Lett* **85**
- Schultz W, Dayan P, Montague PR (1997) **A neural substrate of prediction and reward** *Science* **275**:1593–1599

- Seymour B, O’doherly JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) **Opponent appetitive-aversive neural processes underlie predictive learning of pain relief** *Nat Neurosci* **8**:1234–1240
- Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) **A causal link between prediction errors, dopamine neurons and learning** *Nat Neurosci* **16**:966–973
- Steinmetz NA, Zatzka-Haas P, Carandini M, Harris KD (2019) **Distributed coding of choice, action and engagement across the mouse brain** *Nature* **576**:266–273
- Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) **Dynamic coding for cognitive control in prefrontal cortex** *Neuron* **78**:364–375
- Strait CE, Blanchard TC, Hayden BY (2014) **Reward value comparison via mutual inhibition in ventromedial prefrontal cortex** *Neuron* **82**:1357–1366
- Sutton RS, Barto AG (2018) **Reinforcement learning: An introduction**
- Ten Oever S, Sack AT, Oehrn CR, Axmacher N (2021) **An engram of intentionally forgotten information** *Nat Commun* **12**:1–14
- de Schotten M Thiebaut, Forkel SJ (2022) **The emergent properties of the connected brain** *Science* **378**:505–510
- Thorndike EL (1898) **Animal intelligence: An experimental study of the associative processes in animals** *Psychol Rev Monogr Suppl* **2**
- Tom SM, Fox CR, Trepel C, Poldrack RA (2007) **The Neural Basis of Loss Aversion in Decision-Making Under Risk** *Science* **315**:515–518
- Urai AE, Doiron B, Leifer AM, Churchland AK (2022) **Large-scale neural recordings call for new insights to link brain and behavior** *Nat Neurosci* **25**:11–19
- Varela F, Lachaux J-P, Rodriguez E, Martinerie J (2001) **The brainweb: Phase synchronization and large-scale integration** *Nat Rev Neurosci* **2**:229–239
- Varley TF, Sporns O, Schaffelhofer S, Scherberger H, Dann B (2023) **Information-processing dynamics in neural networks of macaque cerebral cortex reflect cognitive state and behavior** *Proc Natl Acad Sci* **120**
- Vicente R, Wibral M, Lindner M, Pipa G (2011) **Transfer entropy—a model-free measure of effective connectivity for the neurosciences** *J Comput Neurosci* **30**:45–67
- Vidal JR, Chaumon M, O’Regan JK, Tallon-Baudry C (2006) **Visual grouping and the focusing of attention induce gamma-band oscillations at different frequencies in human magnetoencephalogram signals** *J Cogn Neurosci* **18**:1850–1862
- Vinck M, Uran C, Spyropoulos G, Onorato I, Broggin AC, Schneider M, Canales-Johnson A (2023) **Principles of large-scale neural interactions** *Neuron* **111**:987–1002
- Voitov I, Mrsic-Flogel TD (2022) **Cortical feedback loops bind distributed representations of working memory** *Nature* **608**:381–389

Von Siebenthal Z, Boucher O, Rouleau I, Lassonde M, Lepore F, Nguyen DK (2017) **Decision-making impairments following insular and medial temporal lobe resection for drug-resistant epilepsy** *Soc Cogn Affect Neurosci* **12**:128–137

Watkins CJCH, Dayan P (1992) **Q-learning** *Mach Learn* **8**:279–292

Wibral M, Priesemann V, Kay JW, Lizier JT, Phillips WA (2017) **Partial information decomposition as a unified approach to the specification of neural goal functions** *Brain Cogn* **112**:25–38

Wiener N (1956) **The theory of prediction**

Williams PL, Beer RD (2010) **Nonnegative Decomposition of Multivariate Information**

Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C (2006) **Dissociable systems for gain-and loss-related value predictions and errors of prediction in the human brain** *J Neurosci* **26**:9530–9537

Editors

Reviewing Editor

Thorsten Kahnt

National Institute on Drug Abuse Intramural Research Program, Baltimore, United States of America

Senior Editor

Michael Frank

Brown University, Providence, United States of America

Reviewer #1 (Public Review):

Summary:

The work by Combrisson and colleagues investigates the degree to which reward and punishment learning signals overlap in the human brain using intracranial EEG recordings. The authors used information theory approaches to show that local field potential signals in the anterior insula and the three sub regions of the prefrontal cortex encode both reward and punishment prediction errors, albeit to different degrees. Specifically, the authors found that all four regions have electrodes that can selectively encode either the reward or the punishment prediction errors. Additionally, the authors analyzed the neural dynamics across pairs of brain regions and found that the anterior insula to dorsolateral prefrontal cortex neural interactions were specific for punishment prediction errors whereas the ventromedial prefrontal cortex to lateral orbitofrontal cortex interactions were specific to reward prediction errors. This work contributes to the ongoing efforts in both systems neuroscience and learning theory by demonstrating how two differing behavioral signals can be differentiated to a greater extent by analyzing neural interactions between regions as opposed to studying neural signals within one region.

Strengths:

The experimental paradigm incorporates both a reward and punishment component that enables investigating both types of learning in the same group of subjects allowing direct comparisons.

The use of intracranial EEG signals provides much needed insight into the timing of when reward and punishment prediction errors signals emerge in the studied brain regions.

Information theory methods provide important insight into the interregional dynamics associated with reward and punishment learning and allows the authors to assess that reward versus punishment learning can be better dissociated based on interregional dynamics over local activity alone.

Weaknesses:

The analysis presented in the manuscript focuses on gamma band activity. Studying slow oscillations could provide additional insights into the interregional dynamics.

<https://doi.org/10.7554/eLife.92938.2.sa2>

Reviewer #2 (Public Review):

Reward and punishment learning have long been seen as emerging from separate networks of frontal and subcortical areas, often studied separately. Nevertheless, both systems are complimentary and distributed representations of reward and punishments have been repeatedly observed within multiple areas. This raised the unsolved question of the possible mechanisms by which both systems might interact, which this manuscript went after. The authors skillfully leveraged intracranial recordings in epileptic patients performing a probabilistic learning task combined with model-based information theoretical analyses of gamma activities to reveal that information about reward and punishment was not only distributed across multiple prefrontal and insular regions, but that each system showed specific redundant interactions. The reward subsystem was characterized by redundant interactions between orbitofrontal and ventromedial prefrontal cortex, while the punishment subsystem relied on insular and dorsolateral redundant interactions. Finally, the authors revealed a way by which the two systems might interact, through synergistic interaction between ventromedial and dorsolateral prefrontal cortex.

Here, the authors performed an excellent reanalysis of a unique dataset using innovative approaches, pushing our understanding on the interaction at play between prefrontal and insular cortex regions during learning. Importantly, the description of the methods and results is truly made accessible, making it an excellent resource to the community. The authors also carefully report individual subjects' data, which brings confidence in the reproducibility of their observations.

This manuscript goes beyond what is classically performed using intracranial EEG dataset, by not only reporting where a given information, like reward and punishment prediction errors, is represented but also by characterizing the functional interactions that might underlie such representations. The authors highlight the distributed nature of frontal cortex representations and proposed new ways by which the information specifically flows between nodes. This work is well placed to unify our understanding of the complementarity and specificity of the reward and punishment learning systems.

<https://doi.org/10.7554/eLife.92938.2.sa1>

Reviewer #3 (Public Review):

Summary:

The authors investigated that learning processes relied on distinct reward or punishment outcomes in probabilistic instrumental learning tasks were involved in functional interactions of two different cortico-cortical gamma-band modulations, suggesting that learning signals like reward or punishment prediction errors can be processed by two dominated interactions, such as areas IOFC-vmPFC and areas aINS-dlPFC, and later on integrated together in support of switching conditions between reward and punishment learning. By performing the well-known analyses of mutual information, interaction information, and transfer entropy, the conclusion was accomplished by identifying directional task information flow between redundancy-dominated and synergy-dominated interactions. Also, this integral concept provided a unifying view to explain how functional distributed reward and/or punishment information were segregated and integrated across cortical areas.

Strengths:

The dataset used in this manuscript may come from previously published works (Gueguen et al., 2021) or from the same grant project due to the methods. Previous works have shown strong evidence about why gamma-band activities and those 4 areas are important. For further analyses, the current manuscript moved the ideas forward to examine how reward/punishment information transfer between recorded areas corresponding to the task conditions. The standard measurements such mutual information, interaction information, and transfer entropy showed time-series activities in the millisecond level and allowed us to learn the directional information flow during a certain window. In addition, the diagram in Figure 6 summarized the results and proposed an integral concept with functional heterogeneities in cortical areas. These findings in this manuscript will support the ideas from human fMRI studies and add a new insight to electrophysiological studies with the non-human primates.

Comments on revised version:

Thank you authors for all efforts to answer questions from previous comments. I appreciated that authors clarified the terminology and added a paragraph to discuss the current limitations of functional connectivity and anatomical connections. This provided clear and fair explanations to readers who are not familiar with methods in systems neuroscience.

<https://doi.org/10.7554/eLife.92938.2.sa0>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1:

Summary:

The work by Combrisson and colleagues investigates the degree to which reward and punishment learning signals overlap in the human brain using intracranial EEG recordings. The authors used information theory approaches to show that local field potential signals in the anterior insula and the three sub regions of the prefrontal cortex encode both reward and punishment prediction errors, albeit to different degrees. Specifically, the authors found that all four regions have electrodes that can selectively encode either the reward or the punishment prediction errors. Additionally, the authors analyzed the neural dynamics across pairs of brain regions and found that the anterior

insula to dorsolateral prefrontal cortex neural interactions were specific for punishment prediction errors whereas the ventromedial prefrontal cortex to lateral orbitofrontal cortex interactions were specific to reward prediction errors. This work contributes to the ongoing efforts in both systems neuroscience and learning theory by demonstrating how two differing behavioral signals can be differentiated to a greater extent by analyzing neural interactions between regions as opposed to studying neural signals within one region.

Strengths:

The experimental paradigm incorporates both a reward and punishment component that enables investigating both types of learning in the same group of subjects allowing direct comparisons.

The use of intracranial EEG signals provides much needed insight into the timing of when reward and punishment prediction errors signals emerge in the studied brain regions.

Information theory methods provide important insight into the interregional dynamics associated with reward and punishment learning and allows the authors to assess that reward versus punishment learning can be better dissociated based on interregional dynamics over local activity alone.

We thank the reviewer for this accurate summary. Please find below our answers to the weaknesses raised by the reviewer.

Weaknesses:

The analysis presented in the manuscript focuses solely on gamma band activity. The presence and potential relevance of other frequency bands is not discussed. It is possible that slow oscillations, which are thought to be important for coordinating neural activity across brain regions could provide additional insight.

We thank the reviewer for pointing us to this missing discussion in the first version of the manuscript. We now made this point clearer in the Methods sections entitled “iEEG data analysis” and “Estimate of single-trial gamma-band activity”:

“Here, we focused solely on broadband gamma for three main reasons. First, it has been shown that the gamma band activity correlates with both spiking activity and the BOLD fMRI signals (Lachaux et al., 2007; Mukamel et al., 2004; Niessing et al., 2005; Nir et al., 2007), and it is commonly used in MEG and iEEG studies to map task-related brain regions (Brovelli et al., 2005; Crone et al., 2006; Vidal et al., 2006; Ball et al., 2008; Jerbi et al., 2009; Darvas et al., 2010; Lachaux et al., 2012; Cheyne and Ferrari, 2013; Ko et al., 2013). Therefore, focusing on the gamma band facilitates linking our results with the fMRI and spiking literatures on probabilistic learning. Second, single-trial and time-resolved high-gamma activity can be exploited for the analysis of cortico-cortical interactions in humans using MEG and iEEG techniques (Brovelli et al., 2015; 2017; Combrisson et al., 2022). Finally, while previous analyses of the current dataset (Gueguen et al., 2021) reported an encoding of PE signals at different frequency bands, the power in lower frequency bands were shown to carry redundant information compared to the gamma band power.”

The data is averaged across all electrodes which could introduce biases if some subjects had many more electrodes than others. Controlling for this variation in electrode number across subjects would ensure that the results are not driven by a small subset of subjects with more electrodes.

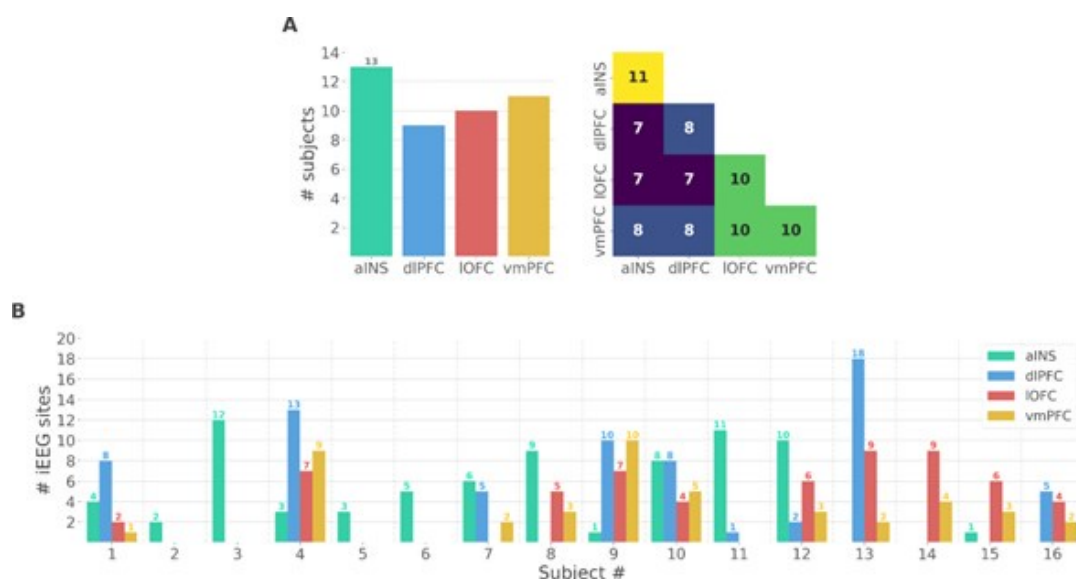
We thank the reviewer for raising this important issue. We would like to point out that the gamma activity was not averaged across bipolar recordings within an area, nor measures of connectivity. Instead, we used a statistical approach proposed in a previous paper that combines non-parametric permutations with measures of information (Combrisson et al., 2022). As we explain in the “Statistical analysis” section, mutual information (MI) is estimated between PE signals and single-trial modulations in gamma activity separately for each contact (or for each pair of contacts). Then, a one-sample t-test is computed across all of the recordings of all subjects to form the effect size at the group-level. We will address the point of the electrode number in our answer below.

The potential variation in reward versus punishment learning across subjects is not included in the manuscript. While the time course of reward versus punishment prediction errors is symmetrical at the group level, it is possible that some subjects show faster learning for one versus the other type which can bias the group average. Subject level behavioral data along with subject level electrode numbers would provide more convincing evidence that the observed effects are not arising from these potential confounds.

We thank the reviewer for the two points raised. We performed additional analyses at the single-participant level to address the issues raised by the reviewer. We should note, however, that these results are descriptive and cannot be generalized to account for population-level effects. As suggested by the reviewer, we prepared two new figures. The first supplementary figure summarizes the number of participants that had iEEG contacts per brain region and pair of brain regions (Fig. S1A in the Appendix). It can be seen that the number of participants sampled in different brain regions is relatively constant (left panel) and the number of participants with pairs of contacts across brain regions is relatively homogeneous, ranging from 7 to 11 (right panel). Fig. S1B shows the number of bipolar derivations per subject and per brain region.

Author response image 1.

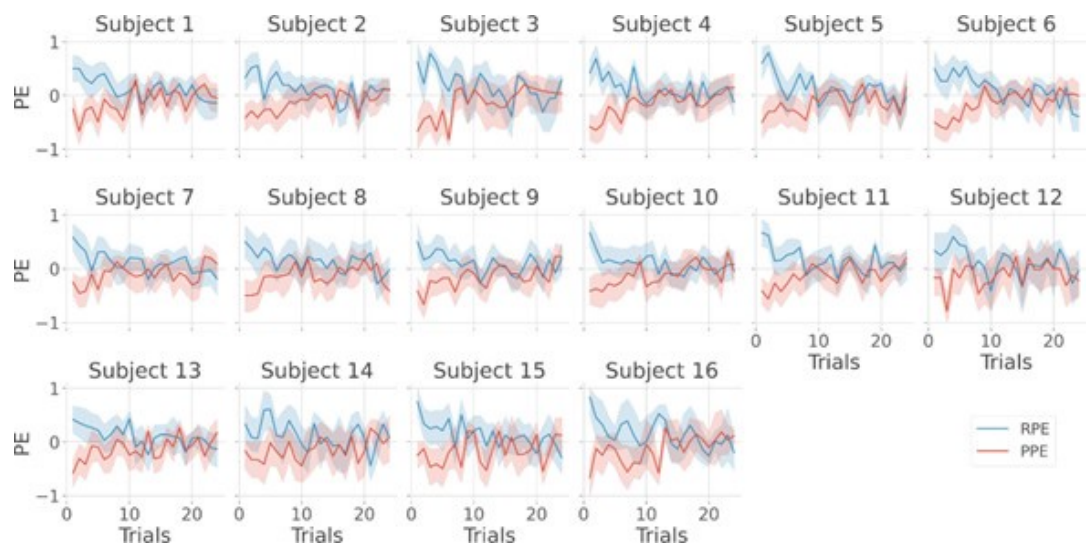
Single subject anatomical repartition. (A) Number of unique subject per brain region and per pair of brain regions (B) Number of bipolar derivations per subject and per brain region



The second supplementary figure describes the estimated prediction error for rewarding and punishing trials for each subject (Fig. S2). The single-subject error bars represent the 95th percentile confidence interval estimated using a bootstrap approach across the different pairs of stimuli presented during the three to six sessions. As the reviewer anticipated, there are indeed variations across subjects, but we observe that RPE and PPE are relatively symmetrical, even at the subject level, and tend toward zero around trial number 10. These results therefore corroborate the patterns observed at the group-level.

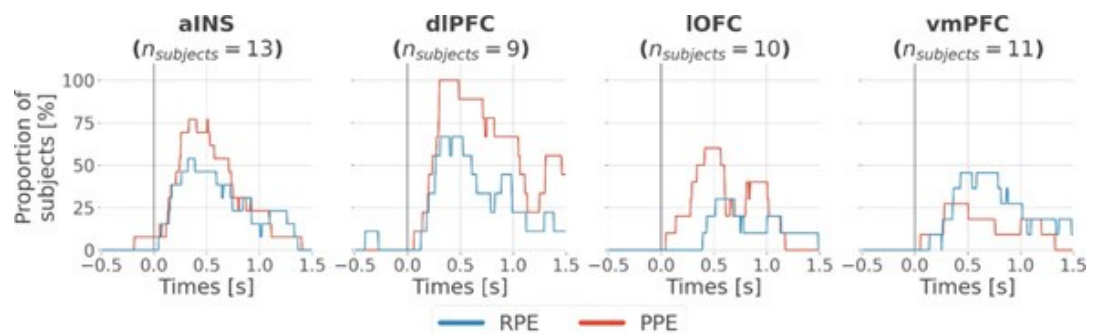
Author response image 2.

Single-subject estimation of predictions errors. Single-subject trial-wise reward PE (RPE - blue) and punishment PE (PPE - red), $\pm$ 95% confidence interval.



Finally, to assess the variability of local encoding of prediction errors across participants, we quantified the proportion of subjects having at least one significant bipolar derivation encoding either the RPE or PPE (Fig. S4). As expected, we found various proportions of unique subjects with significant R/PPE encoding per region. The lowest proportion was achieved in the ventromedial prefrontal cortex (vmPFC) and lateral orbitofrontal cortex (lOFC) for encoding PPE and RPE, respectively, with approximately 30% of the subjects having the effect. Conversely, we found highly reproducible encodings in the anterior insula (aINS) and dorsolateral prefrontal cortex (dlPFC) with a maximum of 100% of the 9 subjects having at least one bipolar derivation encoding PPE in the dlPFC.

Author response image 3.



Taken together, we acknowledge a certain variability per region and per condition. Nevertheless, the results presented in the supplementary figures suggest that the main results do not arise from a minority of subjects.

We would like to point out that in order to assess across-subject variability, a much larger number of participants would have been needed, given the low signal-to-noise ratios observed at the single-participant level. We thus prefer to add these results as supplementary material in the Appendix, rather than in the main text.

It is unclear if the findings in Figures 3 and 4 truly reflect the differential interregional dynamics in reward versus punishment learning or if these results arise as a statistical byproduct of the reward vs punishment bias observed within each region. For instance, the authors show that information transfer from anterior insula to dorsolateral prefrontal cortex is specific to punishment prediction error. However, both anterior insula and dorsolateral prefrontal cortex have higher prevalence of punishment prediction error selective electrodes to begin with. Therefore the findings in Fig 3 may simply be reflecting the prevalence of punishment specificity in these two regions above and beyond a punishment specific neural interaction between the two regions. Either mathematical or analytical evidence that assesses if the interaction effect is simply reflecting the local dynamics would be important to make this result convincing.

This is an important point that we partly addressed in the manuscript. More precisely, we investigated whether the synergistic effects observed between the dlPFC and vmPFC encoding global PEs (Fig. 5) could be explained by their respective local specificity. Indeed, since we reported larger proportions of recordings encoding the PPE in the dlPFC and the RPE in the vmPFC (Fig. 2B), we checked whether the synergy between dlPFC and vmPFC could be mainly due to complementary roles where the dlPFC brings information about the PPE only and the vmPFC brings information to the RPE only. To address this point, we selected PPE-specific bipolar derivations from the dlPFC and RPE-specific from the vmPFC and, as the reviewer predicted, we found synergistic II between the two regions probably mainly because of their respective specificity. In addition, we included the II estimated between non-selective bipolar derivations (i.e. recordings with significant encoding for both RPE and PPE) and we observed synergistic interactions (Fig. 5C and Fig. S9). Taken together, the local specificity certainly plays a role, but this is not the only factor in defining the type of interactions.

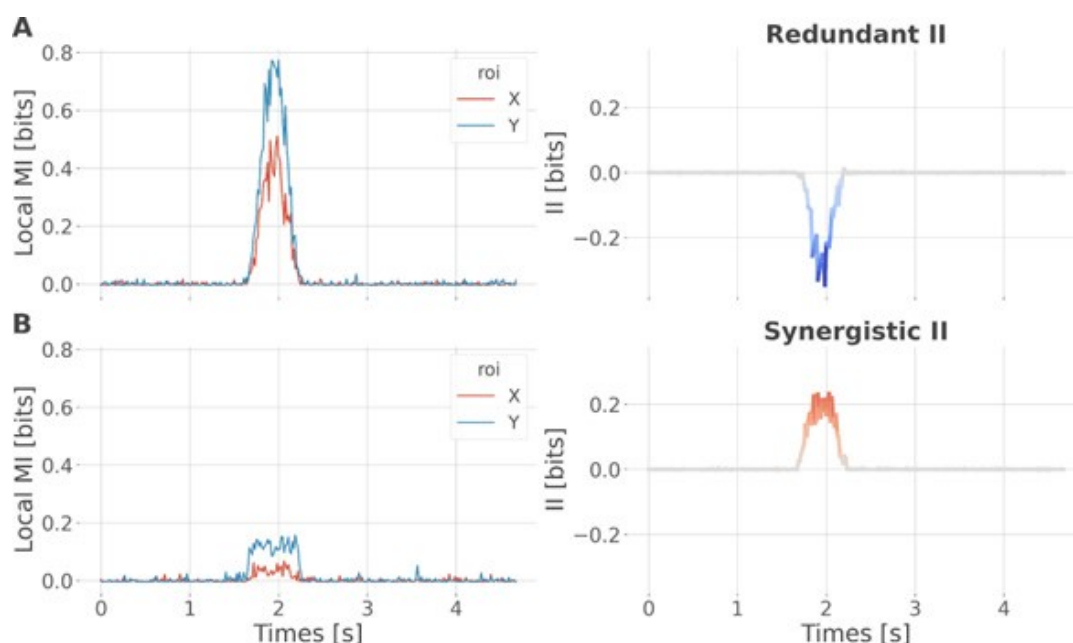
Concerning the interaction information results (II, Fig. 3), several lines of evidence suggest that local specificity cannot account alone for the II effects. For example, the local specificity for PPE is observed across all four areas (Fig. 2A) and the percentage of bipolar derivations displaying an effect is large (equal or above 10%) for three brain regions (aINS, dlPLF and IOFC). If the local specificity were the main driving cause, we would have observed significant redundancy between all pairs of brain regions. On the other hand, the interaction between the aINS and IOFC displayed no significant redundant effect (Fig. 3B). Another example is the result observed in IOFC: approximately 30% of bipolar derivations display a selectivity for PPE (Fig. 2B, third panel from the left), but do not show clear signs of redundant encoding at the level of within-area interactions (Fig. 3A, bottom-left panel). Similarly, the local encoding for RPE is observed across all four brain regions (Fig. 2A) and the percentage of bipolar derivations displaying an effect is large (equal or above 10%) for three brain regions (aINS, dlPLF and vmPFC). Nevertheless, significant between-regions interactions have been observed only between the IOFC and vmPFC (Fig. 3B bottom right panel).

To further support the reasoning, we performed a simulation to show that it is possible to observe synergistic interactions between two regions with the same specificity. As an

example, we may consider one region locally encoding early trials of RPE and a second region encoding the late trials of the RPE. Combining the two with the II would lead to synergistic interactions, because each one of them carries information that is not carried by the other. To illustrate this point, we simulated the data of two regions (x and y). To simulate redundant interactions (first row), each region receives a copy of the prediction (one-to-all) and for the synergy (second row), x and y receive early and late PE trials, respectively (all-to-one). This toy example illustrates that the local specificity is not the only factor determining the type of their interactions. We added the following result to the Appendix.

Author response image 4.

Local specificity does not fully determine the type of interactions. Within-area local encoding of PE using the mutual information (MI, in bits) for regions X and Y and between-area interaction information (II, in bits) leading to (A) redundant interactions and (B) synergistic interactions about the PE



Regarding the information transfer results (Fig. 4), similar arguments hold and suggest that the prevalence is not the main factor explaining the arising transfer entropy between the anterior insula (aINS) and dorsolateral prefrontal cortex (dlPFC). Indeed, the IOFC has a strong local specificity for PPE, but the transfer entropy between the IOFC and aINS (or dlPFC) is shown in Fig. S7 does not show significant differences in encoding between PPE and RPE.

Indeed, such transfer can only be found when there is a delay between the gamma activity of the two regions. In this example, the transfer entropy quantifies the amount of information shared between the past activity of the aINS and the present activity of the dlPFC conditioned on the past activity of the dlPFC. The conditioning ensures that the present activity of the dlPFC is not only explained by its own past. Consequently, if both regions exhibit various prevalences toward reward and punishment but without delay (i.e. at the same timing), the transfer entropy would be null because of the conditioning. As a fact, between 10 to -20% of bipolar recordings show a selectivity to the reward PE (represented by a proportion of 40-60% of subjects, Fig.S4). However, the transfer entropy estimated from the aINS to the dlPFC across rewarding trials is flat and clearly non-significant. If the transfer entropy was a byproduct of the local specificity then we should observe an increase, which is not the case here.

Reviewer #2:

Summary:

Reward and punishment learning have long been seen as emerging from separate networks of frontal and subcortical areas, often studied separately. Nevertheless, both systems are complimentary and distributed representations of rewards and punishments have been repeatedly observed within multiple areas. This raised the unsolved question of the possible mechanisms by which both systems might interact, which this manuscript went after. The authors skillfully leveraged intracranial recordings in epileptic patients performing a probabilistic learning task combined with model-based information theoretical analyses of gamma activities to reveal that information about reward and punishment was not only distributed across multiple prefrontal and insular regions, but that each system showed specific redundant interactions. The reward subsystem was characterized by redundant interactions between orbitofrontal and ventromedial prefrontal cortex, while the punishment subsystem relied on insular and dorsolateral redundant interactions. Finally, the authors revealed a way by which the two systems might interact, through synergistic interaction between ventromedial and dorsolateral prefrontal cortex.

Strengths:

Here, the authors performed an excellent reanalysis of a unique dataset using innovative approaches, pushing our understanding on the interaction at play between prefrontal and insular cortex regions during learning. Importantly, the description of the methods and results is truly made accessible, making it an excellent resource to the community.

This manuscript goes beyond what is classically performed using intracranial EEG dataset, by not only reporting where a given information, like reward and punishment prediction errors, is represented but also by characterizing the functional interactions that might underlie such representations. The authors highlight the distributed nature of frontal cortex representations and propose new ways by which the information specifically flows between nodes. This work is well placed to unify our understanding of the complementarity and specificity of the reward and punishment learning systems.

We thank the reviewer for the positive feedback. Please find below our answers to the weaknesses raised by the reviewer.

Weaknesses:

The conclusions of this paper are mostly supported by the data, but whether the findings are entirely generalizable would require further information/analyses.

First, the authors found that prediction errors very quickly converge toward 0 (less than 10 trials) while subjects performed the task for sets of 96 trials. Considering all trials, and therefore having a non-uniform distribution of prediction errors, could potentially bias the various estimates the authors are extracting. Separating trials between learning (at the start of a set) and exploiting periods could prove that the observed functional interactions are specific to the learning stages, which would strengthen the results.

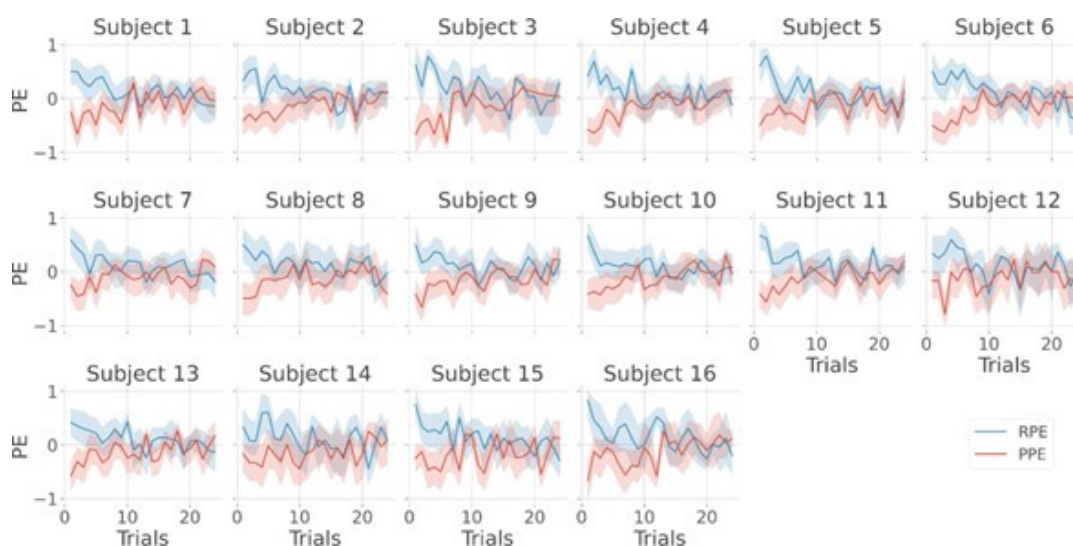
We thank the reviewer for this question. We would like to note that the probabilistic nature of the learning task does not allow a strict distinction between the exploration and exploitation phases. Indeed, the probability of obtaining the less rewarding outcome was 25% (i.e., for 0€ gain in the reward learning condition and -1€ loss in the punishment learning condition). Thus, participants tended to explore even during the last set of trials in each

session. This is evident from the average learning curves shown in Fig. 1B of (Gueguen et al., 2021). Learning curves show rates of correct choice (75% chance of 1€ gain) in the reward condition (blue curves) and incorrect choice (75% chance of 1€ loss) in the punishment condition (red curves).

For what concerns the evolution of PEs, as reviewer #1 suggested, we added a new figure representing the single-subject estimates of the R/PPE (Fig S2). Here, the confidence interval is obtained across all pairs of stimuli presented during the different sessions. We retrieved the general trend of the R/PPE converging toward zero around 10 trials. Both average reward and punishment prediction errors converge toward zero in approximately 10 trials, single-participant curves display large variability, also at the end of each session. As a reminder, the 96 trials represent the total number of trials for one session for the four pairs and the number of trials for each stimulus was only 24.

Author response image 5.

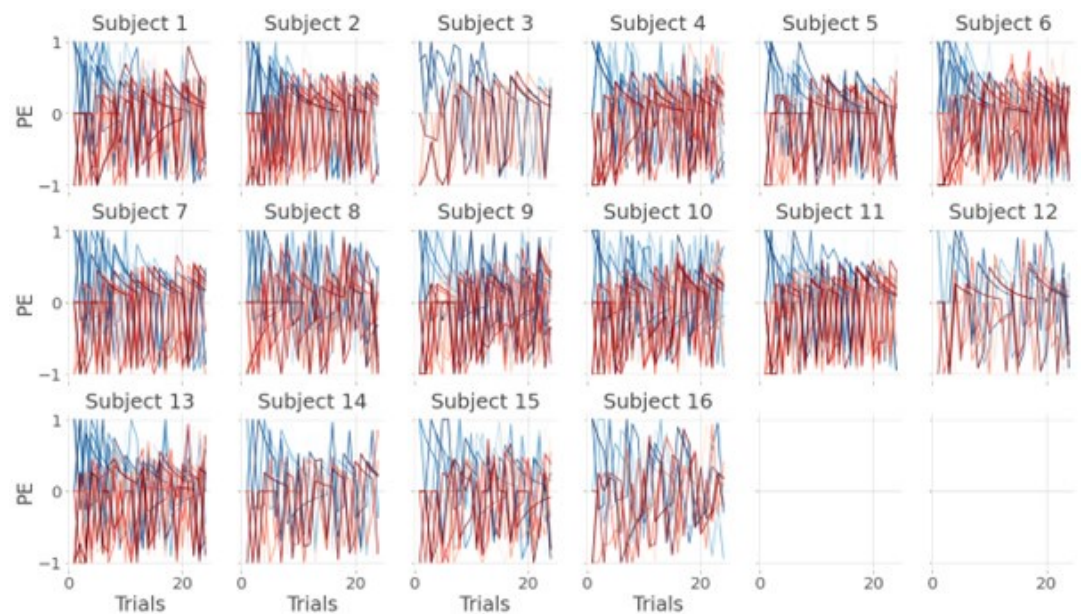
Single-subject estimation of predictions errors. Single-subject trial-wise reward PE (RPE - blue) and punishment PE (PPE - red), $\pm$ 95% confidence interval



However, the convergence of the R/PPE is due to the average across the pairs of stimuli. In the figure below, we superimposed the estimated R/PPE, per pair of stimuli, for each subject. It becomes very clear that high values of PE can be reached, even for late trials. Therefore, we believe that the split into early/late trials because of the convergence of PE is far from being trivial.

Author response image 6.

Single-subject estimation of predictions errors per pair of stimuli. Single-subject trial-wise reward PE (RPE - blue) and punishment PE (PPE - red)

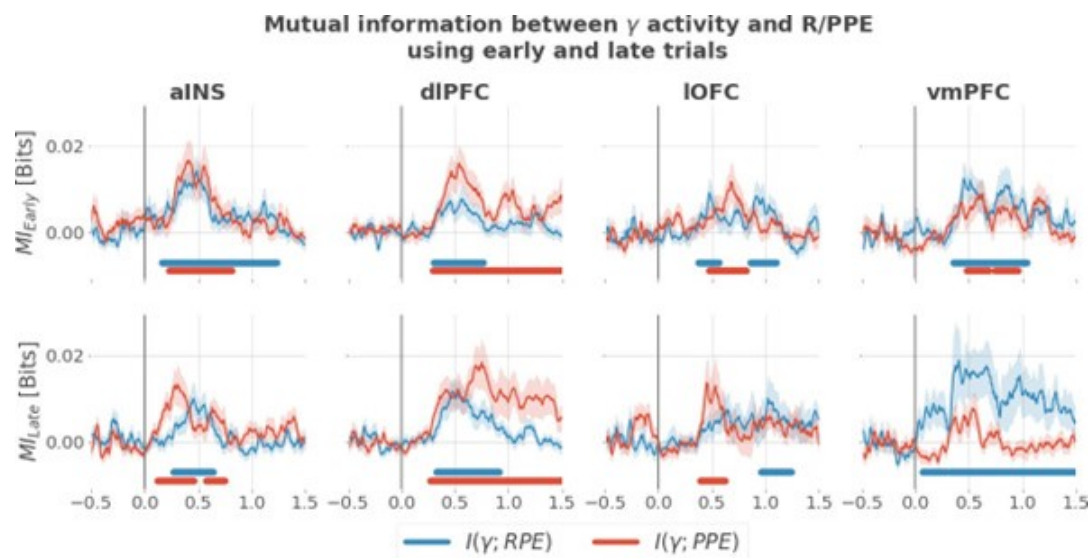


Consequently, nonzero PRE and PPE occur during the whole session and separating trials between learning (at the start of a set) and exploiting periods, as suggested by the reviewer, does not allow a strict dissociation between learning vs no-learning. Nevertheless, we tested the analysis proposed by the reviewer, at the local level. We splitted the 24 trials of each pair of stimuli into early, middle and late trials (8 trials each). We then reproduced Fig. 2 by computing the mutual information between the gamma activity and the R/PPE for subsets of trials: early (first row) and late trials (second row). We retrieved significant encoding of both R/PPE in the aINS, dlPFC and lOFC in both early and late trials. The vmPFC also showed significant encoding of both during early trials. The only difference emerges in the late trials of the vmPFC where we found a strong encoding of the RPE only. It should also be noted that here since we are sub-selecting the trials, the statistical analyses are only performed using a third of the trials.

Taken together, the combination of high values of PE achieved even for late trials and the fact that most of the findings are reproduced even with a third of the trials does not justify the split into early and late trials here. Crucially, this latest analysis confirms that the neural correlates of learning that we observed reflect PE signals rather than early versus late trials in the session.

Author response image 7.

MI between gamma activity and R/PPE using early and late trials. Time courses of MI estimated between the gamma power and both RPE (blue) and PPE (red) using either early or late trials (first and second row, respectively). Horizontal thick lines represent significant clusters of information ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs).

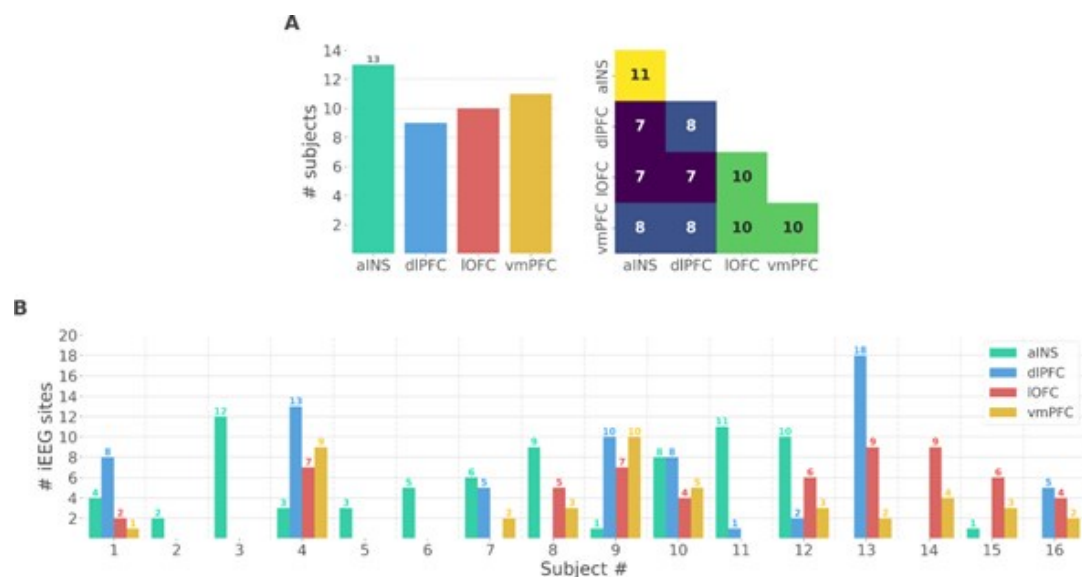


Importantly, it is unclear whether the results described are a common feature observed across subjects or the results of a minority of them. The authors should report and assess the reliability of each result across subjects. For example, the authors found RPE-specific interactions between vmPFC and IOFC, even though less than 10% of sites represent RPE or both RPE/PPE in IOFC. It is questionable whether such a low proportion of sites might come from different subjects, and therefore whether the interactions observed are truly observed in multiple subjects. The nature of the dataset obviously precludes from requiring all subjects to show all effects (given the known limits inherent to intracerebral recording in patients), but it should be proven that the effects were reproducibly seen across multiple subjects.

We thank the reviewer for this remark that has also been raised by the first reviewer. This issue was raised by the first reviewer. Indeed, we added a supplementary figure describing the number of unique subjects per brain region and per pair of brain regions (Fig. S1A) such as the number of bipolar derivations per region and per subject (Fig. S1B).

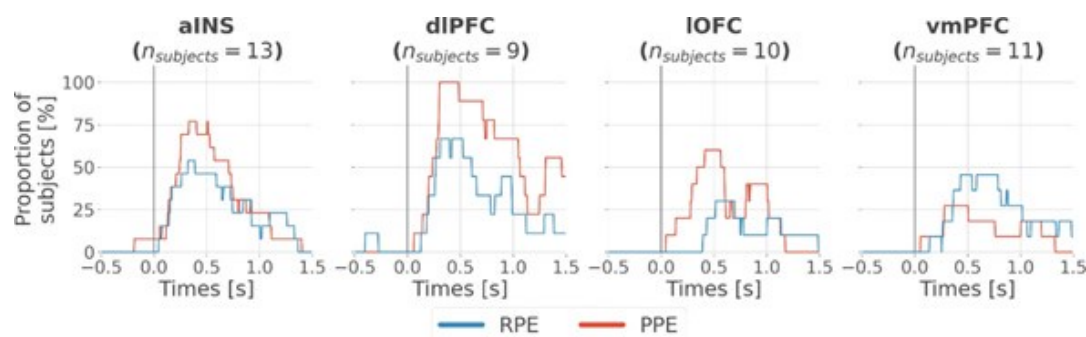
Author response image 8.

Single subject anatomical repartition. (A) Number of unique subject per brain region and per pair of brain regions (B) Number of bipolar derivations per subject and per brain region



Regarding the reproducibility of the results across subjects for the local analysis (Fig. 2), we also added the instantaneous proportion of subjects having at least one bipolar derivation showing a significant encoding of the RPE and PPE (Fig. S4). We found a minimum proportion of approximately 30% of unique subjects having the effect in the IOFC and vmPFC, respectively with the RPE and PPE. On the other hand, both the aINS and dlPFC showed between 50 to 100% of the subjects having the effect. Therefore, local encoding of RPE and PPE was never represented by a single subject.

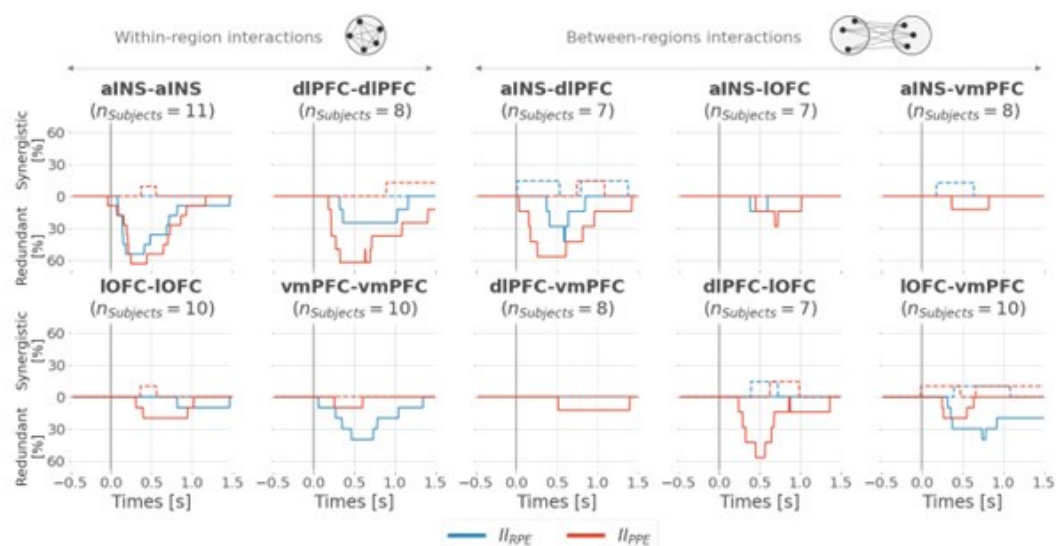
Author response image 9.



Similarly, we performed statistical analysis on interaction information at the single-subject level and counted the proportion of unique subjects having at least one pair of recordings with significant redundant and synergistic interactions about the RPE and PPE (Fig. S5). Consistently with the results shown in Fig. 3, the proportions of significant redundant and synergistic interactions are negative and positive, respectively. For the within-regions interactions, approximately 60% of the subjects with redundant interactions are about R/PPE in the aINS and about the PPE in the dlPFC and 40% about the RPE in the vmPFC. For the across-regions interactions, 60% of the subjects have redundant interactions between the aINS-dlPFC and dlPFC-IOFC about the PPE, and 30% have redundant interactions between IOFC-vmPFC about the RPE. Globally, we reproduced the main results shown in Fig. 3.

Author response image 10.

Inter-subjects reproducibility of redundant interactions about PE signals. Time-courses of proportion of subjects having at least one pair of bipolar derivation with a significant interaction information ($p < 0.05$, cluster-based correction, non-parametric randomization across epochs) about the RPE (blue) or PPE (red). Data are aligned to the outcome presentation (vertical line at 0 seconds). Proportion of subjects with redundant (solid) and synergistic (dashed) interactions are respectively going downward and upward.



Finally, the timings of the observed interactions between areas preclude one of the authors' main conclusions. Specifically, the authors repeatedly concluded that the encoding of RPE/PPE signals are "emerging" from redundancy-dominated prefrontal-insular interactions. However, the between-region information and transfer entropy between vmPFC and IOFC for example is observed almost 500ms after the encoding of RPE/PPE in these regions, questioning how it could possibly lead to the encoding of RPE/PPE. It is also noteworthy that the two information measures, interaction information and transfer entropy, between these areas happened at non overlapping time windows, questioning the underlying mechanism of the communication at play (see Figures 3/4). As an aside, when assessing the direction of information flow, the authors also found delays between pairs of signals peaking at 176ms, far beyond what would be expected for direct communication between nodes. Discussing this aspect might also be of importance as it raises the possibility of third-party involvement.

The local encoding of RPE in the vmPFC and IOFC is observed in a time interval ranging from approximately 0.2-0.4s to 1.2-1.4s after outcome presentation (blue bars in **Fig. 2A**). The encoding of RPE by interaction information covers a time interval from approximately 1.1s to 1.5s (blue bars in **Fig. 3B**, bottom right panel). Similarly, significant TE modulations between the vmPFC and IOFC specific for PPE occur mainly in the 0.7s-1.1s range. Thus, it seems that the local encoding of PPE precedes the effects observed at the level of the neural interactions (II and TE). On the other hand, the modulations in MI, II and TE related to PPE co-occur in a time window from 0.2s to 0.7s after outcome presentation. Thus, we agree with the reviewer that a generic conclusion about the potential mechanisms relating the three levels of analysis cannot be drawn. We thus replaced the term "emerge from" by "occur with" from the manuscript which may be misinterpreted as hinting at a potential mechanism. We nevertheless concluded that the three levels of analysis (and phenomena) co-occur in time,

thus hinting at a potential across-scales interaction that needs further study. Indeed, our study suggests that further work, beyond the scope of the current study, is required to better understand the interaction between scales.

Regarding the delay for the conditioning of the transfer entropy, the value of 176 ms reflects the delay at which we observed a maximum of transfer entropy. However, we did not use a single delay for conditioning, we used every possible delay between [116, 236] ms, as explained in the Method section. We would like to stress that transfer entropy is a directed metric of functional connectivity, and it can only be interpreted as quantifying statistical causality defined in terms of predictability according to the Wiener-Granger principle, as detailed in the methods. Thus, it cannot be interpreted in Pearl's causal terms and as indexing any type of direct communication between nodes. This is a known limitation of the method, which has been stressed in past literature and that we believe does not need to be addressed here.

To account for this, we revised the discussion to make sure this issue is addressed in the following paragraph:

“Here, we quantified directional relationships between regions using the transfer entropy (Schreiber, 2000), which is a functional connectivity measure based on the Granger-Wiener causality principle. Tract tracing studies in the macaque have revealed strong interconnections between the IOFC and vmPFC in the macaque (Carmichael and Price, 1996; Öngür and Price, 2000). In humans, cortico-cortical anatomical connections have mainly been investigated using diffusion magnetic resonance imaging (dMRI). Several studies found strong probabilities of structural connectivity between the anterior insula with the orbitofrontal cortex and dorsolateral part of the prefrontal cortex (Cloutman et al., 2012; Ghaziri et al., 2017), and between the IOFC and vmPFC (Heather Hsu et al., 2020). In addition, the statistical dependency (e.g. coherence) between the LFP of distant areas could be potentially explained by direct anatomical connections (Schneider et al., 2021; Vinck et al., 2023). Taken together, the existence of an information transfer might rely on both direct or indirect structural connectivity. However, here we also reported differences of TE between rewarding and punishing trials given the same backbone anatomical connectivity (Fig. 4). [...]”

Reviewer #3:

Summary:

The authors investigated that learning processes relied on distinct reward or punishment outcomes in probabilistic instrumental learning tasks were involved in functional interactions of two different cortico-cortical gamma-band modulations, suggesting that learning signals like reward or punishment prediction errors can be processed by two dominated interactions, such as areas IOFC-vmPFC and areas aINS-dIPFC, and later on integrated together in support of switching conditions between reward and punishment learning. By performing the well-known analyses of mutual information, interaction information, and transfer entropy, the conclusion was accomplished by identifying directional task information flow between redundancy-dominated and synergy-dominated interactions. Also, this integral concept provided a unifying view to explain how functional distributed reward and/or punishment information were segregated and integrated across cortical areas.

Strengths:

The dataset used in this manuscript may come from previously published works (Gueguen et al., 2021) or from the same grant project due to the methods. Previous works have shown strong evidence about why gamma-band activities and those 4 areas

are important. For further analyses, the current manuscript moved the ideas forward to examine how reward/punishment information transfer between recorded areas corresponding to the task conditions. The standard measurements such as mutual information, interaction information, and transfer entropy showed time-series activities in the millisecond level and allowed us to learn the directional information flow during a certain window. In addition, the diagram in Figure 6 summarized the results and proposed an integral concept with functional heterogeneities in cortical areas. These findings in this manuscript will support the ideas from human fMRI studies and add a new insight to electrophysiological studies with the non-human primates.

We thank the reviewer for the summary such as for highlighting the strengths. Please find below our answers regarding the weaknesses of the manuscript.

Weaknesses:

After reading through the manuscript, the term "non-selective" in the abstract confused me and I did not actually know what it meant and how it fits the conclusion. If I learned the methods correctly, the 4 areas were studied in this manuscript because of their selective responses to the RPE and PPE signals (Figure 2). The redundancy- and synergy-dominated subsystems indicated that two areas shared similar and complementary information, respectively, due to the negative and positive value of interaction information (Page 6). For me, it doesn't mean they are "non-selective", especially in redundancy-dominated subsystem. I may miss something about how you calculate the mutual information or interaction information. Could you elaborate this and explain what the "non-selective" means?

In the study performed by Gueguen et al. in 2021, the authors used a general linear model (GLM) to link the gamma activity to both the reward and punishment prediction errors and they looked for differences between the two conditions. Here, we reproduced this analysis except that we used measures from the information theory (mutual information) that were able to capture linear and non-linear relationships (although monotonic) between the gamma activity and the prediction errors. The clusters we reported reflect significant encoding of either the RPE and/or the PPE. From **Fig. 2**, it can be seen that the four regions have a gamma activity that is modulated according to both reward and punishment PE. We used the term “non-selective”, because the regions did not encode either one or the other, but various proportions of bipolar derivations encoding either one or both of them.

The directional information flows identified in this manuscript were evidenced by the recording contacts of iEEG with levels of concurrent neural activities to the task conditions. However, are the conclusions well supported by the anatomical connections? Is it possible that the information was transferred to the target via another area? These questions may remain to be elucidated by using other approaches or animal models. It would be great to point this out here for further investigation.

We thank the reviewer for this interesting question. We added the following paragraph to the discussion to clarify the current limitations of the transfer entropy and the link with anatomical connections :

“Here, we quantified directional relationships between regions using the transfer entropy (Schreiber, 2000), which is a functional connectivity measure based on the Granger-Wiener causality principle. Tract tracing studies in the macaque have revealed strong interconnections between the IOFC and vmPFC in the macaque (Carmichael and Price, 1996; Öngür and Price, 2000). In humans, cortico-cortical anatomical connections have mainly been investigated using diffusion magnetic resonance imaging (dMRI). Several studies found strong probabilities of structural connectivity between the anterior insula with the

orbitofrontal cortex and dorsolateral part of the prefrontal cortex (Cloutman et al., 2012; Ghaziri et al., 2017), and between the IOFC and vmPFC (Heather Hsu et al., 2020). In addition, the statistical dependency (e.g. coherence) between the LFP of distant areas could be potentially explained by direct anatomical connections (Schneider et al., 2021). Taken together, the existence of an information transfer might rely on both direct or indirect structural connectivity. However, here we also reported differences of TE between rewarding and punishing trials given the same backbone anatomical connectivity (Fig. 4). Our results are further supported by a recent study involving drug-resistant epileptic patients with resected insula who showed poorer performance than healthy controls in case of risky loss compared to risky gains (Von Siebenthal et al., 2017).”

References

- Carmichael ST, Price J. 1996. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol* 371:179–207.
- Cloutman LL, Binney RJ, Drakesmith M, Parker GJM, Lambon Ralph MA. 2012. The variation of function across the human insula mirrors its patterns of structural connectivity: Evidence from in vivo probabilistic tractography. *NeuroImage* 59:3514–3521. doi:10.1016/j.neuroimage.2011.11.016
- Combrisson E, Allegra M, Basanisi R, Ince RAA, Giordano BL, Bastin J, Brovelli A. 2022. Group-level inference of information-based measures for the analyses of cognitive brain networks from neurophysiological data. *NeuroImage* 258:119347. doi:10.1016/j.neuroimage.2022.119347
- Ghaziri J, Tucholka A, Girard G, Houde J-C, Boucher O, Gilbert G, Descoteaux M, Lippé S, Rainville P, Nguyen DK. 2017. The Corticocortical Structural Connectivity of the Human Insula. *Cereb Cortex* 27:1216–1228. doi:10.1093/cercor/bhv308
- Gueguen MCM, Lopez-Persem A, Billeke P, Lachaux J-P, Rheims S, Kahane P, Minotti L, David O, Pessiglione M, Bastin J. 2021. Anatomical dissociation of intracerebral signals for reward and punishment prediction errors in humans. *Nat Commun* 12:3344. doi:10.1038/s41467-021-23704-w
- Heather Hsu C-C, Rolls ET, Huang C-C, Chong ST, Zac Lo C-Y, Feng J, Lin C-P. 2020. Connections of the Human Orbitofrontal Cortex and Inferior Frontal Gyrus. *Cereb Cortex* 30:5830–5843. doi:10.1093/cercor/bhaa160
- Lachaux J-P, Fonlupt P, Kahane P, Minotti L, Hoffmann D, Bertrand O, Baciau M. 2007. Relationship between task-related gamma oscillations and BOLD signal: new insights from combined fMRI and intracranial EEG. *Hum Brain Mapp* 28:1368–1375. doi:10.1002/hbm.20352
- Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. 2004. Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex. *Cereb Cortex* 14:881.
- Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA. 2005. Hemodynamic signals correlate tightly with synchronized gamma oscillations. *science* 309:948–951.
- Nir Y, Fisch L, Mukamel R, Gelbard-Sagiv H, Arieli A, Fried I, Malach R. 2007. Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr Biol* 17:1275–1285.
- Öngür D, Price JL. 2000. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10:206–219.
- Schneider M, Brogini AC, Dann B, Tzanou A, Uran C, Sheshadri S, Scherberger H, Vinck M. 2021. A mechanism for inter-areal coherence through communication based on connectivity

and oscillatory power. *Neuron* 109:4050-4067.e12. doi:10.1016/j.neuron.2021.09.037

Schreiber T. 2000. Measuring information transfer. *Phys Rev Lett* 85:461.

Von Siebenthal Z, Boucher O, Rouleau I, Lassonde M, Lepore F, Nguyen DK. 2017. Decision-making impairments following insular and medial temporal lobe resection for drug-resistant epilepsy. *Soc Cogn Affect Neurosci* 12:128–137. doi:10.1093/scan/nsw152

Recommendations for the authors

Reviewer #1

(1) Overall, the writing of the manuscript is dense and makes it hard to follow the scientific logic and appreciate the key findings of the manuscript. I believe the manuscript would be accessible to a broader audience if the authors improved the writing and provided greater detail for their scientific questions, choice of analysis, and an explanation of their results in simpler terms.

We extensively modified the introduction to better describe the rationale and research question.

(2) In the introduction the authors state "we hypothesized that reward and punishment learning arise from complementary neural interactions between frontal cortex regions". This stated hypothesis arrives rather abruptly after a summary of the literature given that the literature summary does not directly inform their stated hypothesis. Put differently, the authors should explicitly state what the contradictions and/or gaps in the literature are, and what specific combinations of findings guide them to their hypothesis. When the authors state their hypothesis the reader is still left asking: why are the authors focusing on the frontal regions? What do the authors mean by complementary interactions? What specific evidence or contradiction in the literature led them to hypothesize that complementary interactions between frontal regions underlie reward and punishment learning?

We extensively modified the introduction and provided a clearer description of the brain circuits involved and the rationale for searching redundant and synergistic interactions between areas.

(3) Related to the above point: when the authors subsequently state "we tested whether redundancy- or synergy dominated interactions allow the emergence of collective brain networks differentially supporting reward and punishment learning", the Introduction (up to the point of this sentence) has not been written to explain the synergy vs. redundancy framework in the literature and how this framework comes into play to inform the authors' hypothesis on reward and punishment learning.

We extensively modified the introduction and provided a clearer description of redundant and synergistic interactions between areas.

(4) The explanation of redundancy vs synergy dominated brain networks itself is written densely and hard to follow. Furthermore, how this framework informs the question on the neural substrates of reward versus punishment learning is unclear. The authors should provide more precise statements on how and why redundancy vs. synergy comes into play in reward and punishment learning. Put differently, this redundancy vs. synergy framework is key for understanding the manuscript and the introduction is not written clearly enough to explain the framework and how it informs the authors' hypothesis and research questions on the neural substrates of reward vs. punishment learning.

Same as above

(5) While the choice of these four brain regions in context of reward and punishment learning does make sense, the authors do not outline a clear scientific justification as to why these regions were selected in relation to their question.

Same as above

(6) Could the authors explain why they used gamma band power (as opposed to or in addition to the lower frequency bands) to investigate MI. Relatedly, when the authors introduce MI analysis, it would be helpful to briefly explain what this analysis measures and why it is relevant to address the question they are asking.

Please see our answer to the first public comment. We added a paragraph to the discussion section to justify our choice of focusing on the gamma band only. We added the following sentence to the result section to justify our choice for using mutual-information:

The MI allowed us to detect both linear and non-linear relationships between the gamma activity and the PE

An extended explanation justifying our choice for the MI was already present in the method section.

(7) The authors state that "all regions displayed a local "probabilistic" encoding of prediction errors with temporal dynamics peaking around 500 ms after outcome presentation". It would be helpful for the reader if the authors spelled out what they mean by probabilistic in this context as the term can be interpreted in many different ways.

We agree with the reviewer that the term “probabilistic” can be interpreted in different ways. In the revised manuscript we changed “probabilistic” for “mixed”.

(8) The authors should include a brief description of how they compute RPE and PPE in the beginning of the relevant results section.

The explanation of how we estimated the PE is already present in the result section: “We estimated trial-wise prediction errors by fitting a Q-learning model to behavioral data. Fitting the model consisted in adjusting the constant parameters to maximize the likelihood of observed choices etc.”

(9) It is unclear from the Methods whether the authors have taken any measures to address the likely difference in the number of electrodes across subjects. For example, it is likely that some subjects have 10 electrodes in vmPFC while others may have 20. In group analyses, if the data is simply averaged across all electrodes then each subject contributes a different number of data points to the analysis. Hence, a subject with more electrodes can bias the group average. A starting point would be to state the variation in number of electrodes across subjects per brain region. If this variation is rather small, then simple averaging across electrodes might be justified. If the variation is large then one idea would be to average data across electrodes within subjects prior to taking the group average or use a resampling approach where the minimum number of electrodes per brain area is subsampled.

We addressed this point in our public answers. As a reminder, the new version of the manuscript contains a figure showing the number of unique patients per region, the PE at per

participant level together with local-encoding at the single participant level.

(10) One thing to consider is whether the reward and punishment in the task is symmetrical in valence. While 1\$ increase and 1\$ decrease is equivalent in magnitude, the psychological effect of the positive (vs. the negative) outcome may still be asymmetrical and the direction and magnitude of this asymmetry can vary across individuals. For instance, some subjects may be more sensitive to the reward (over punishment) while others are more sensitive to the punishment (over reward). In this scenario, it is possible that the differentiation observed in PPE versus RPE signals may arise from such psychological asymmetry rather than the intrinsic differences in how certain brain regions (and their interactions) may encode for reward vs punishment. Perhaps the authors can comment on this possibility, and/or conduct more in depth behavioral analysis to determine if certain subjects adjust their choice behavior faster in response to reward vs. punishment contexts.

While it could be possible that individuals display different sensitivities vis-à-vis positive and negative prediction errors (and, indeed, a vast body of human reinforcement learning literature seems to point in this direction; Palminteri & Lebreton, 2022), it is unclear to us how such differences would explain into the recruitment of anatomically distinct areas reward and punishment prediction errors. It is important to note here that our design partially orthogonalized positive and reward vs. negative and punishment PEs, because the neutral outcome can generate both positive and negative prediction errors, as a function of the learning context (reward-seeking and punishment avoidance). Back to the main question, for instance, Lefebvre et al (2017) investigated with fMRI the neural correlates of reward prediction errors only and found that inter-individual differences in learning rates for positive and negative prediction errors correlated with differences in the degree of striatal activation and not with the recruitment of different areas. To sum up, while we acknowledge that individuals may display different sensitivity to prediction errors (and reward magnitudes), we believe that such differences should translated in difference in the degree of activation of a given system (the reward systems vs the punishment one) rather than difference in neural system recruitment

(11) As summarized in Fig 6, the authors show that information transfer between aINS to dlPFC was PPE specific whereas the information transfer between vmPFC to IOFC was RPE specific. What is unclear is if these findings arise as an inevitable statistical byproduct of the fact that aINS has high PPE-specificity and that vmPFC has high RPE-specificity. In other words, it is possible that the analysis in Fig 3,4 are sensitive to fact that there is a larger proportion of electrodes with either PPE or RPE sensitivity in aINS and vmPFC respectively - and as such, the II analysis might reflect the dominant local encoding properties above and beyond reflecting the interactions between regions per se. Simply put, could the analysis in Fig 3B turn out in any other way given that there are more PPE specific electrodes in aINS and more RPE specific electrodes in vmPFC? Some options to address this question would be to limit the electrodes included in the analyses (in Fig 3B for example) so that each region has the same number of PPE and RPE specific electrodes included.

Please see the simulation we added to the revised manuscript (Fig. S10) demonstrating that synergistic interactions can emerge between regions with the same specificity.

Regarding the possibility that Fig. 3 and 4 are sensitive to the number of bipolar derivations being R/PPE specific, a counter-example is the vmPFC. The vmPFC has a few recordings specific to punishment (Fig. 2) in almost 30% of the subjects (Fig. S4). However, there is no II about the PPE between recordings of the vmPFC (Fig. 3). The same reasoning also holds for

the IOFC. Therefore, the proportion of recordings being RPE or PPE-specific is not sufficient to determine the type of interactions.

(12) Related to the point above, what would the results presented in Fig 3A (and 3B) look like if the authors ran the analyses on RPE specific and PPE specific electrodes only. Is the vmPFC-vmPFC RPE effect in Fig 3A arising simply due to the high prevalence of RPE specific electrodes in vmPFC (as shown in Fig. 2)?

Please see our answer above.

Reviewer #2:

Regarding Figure 2A, the authors argued that their findings "globally reproduced their previously published findings" (from Gueguen et al, 2021). It is worth noting though that in their original analysis, both aINS and IOFC show differential effects (aINS showing greater punishment compared to reward, and the opposite for IOFC) compared to the current analysis. Although I would be akin to believe that the nonlinear approach used here might explain part of the differences (as the authors discussed), I am very wary of the other argument advanced: "the removal of iEEG sites contaminated with pathological activity". This raised some red flags. Does that mean some of the conclusions observed in Gueguen et al (2021) are only the result of noise contamination, and therefore should be disregarded? The author might want to add a short supplementary figure using the same approach as in Gueguen (2021) but using the subset of contacts used here to comfort potential readers of the validity of their previous manuscript.

We appreciate the reviewer's concerns and understand the request for additional information. However, we would like to point out that the figure suggested by the reviewer is already present in the supplementary files of Gueguen et al. 2021 (see Fig. S2). The results of this study should not be disregarded, as the supplementary figure reproduces the results of the main text after excluding sites with pathological activity. Including or excluding sites contaminated with epileptic activity does not have a significant impact on the results, as analyses are performed at each time-stamp and across trials, and epileptic spikes are never aligned in time across trials.

That being said, there are some methodological differences between the two studies. To extract gamma power, Gueguen et al. filtered and averaged 10 Hz sub-bands, while we used multi-tapers. Additionally, they used a temporal smoothing of 250 ms, while we used less smoothing. However, as explained in the main text, we used information-theoretical approaches to capture the statistical dependencies between gamma power and PE. Despite divergent methodologies, we obtained almost identical results.

The data and code supporting this manuscript should be made available. If raw data cannot be shared for ethical reasons, single-trial gamma activities should at least be provided. Regarding the code used to process the data, sharing it could increase the appeal (and use) of the methods applied.

We thank the reviewer for this suggestion. We added a section entitled "Code and data availability" and gave links to the scripts, notebooks and preprocessed data.

<https://doi.org/10.7554/eLife.92938.2.sa4>