# Higher-order and distributed synergistic functional interactions encode information gain in goal-directed learning

Etienne Combrisson[1]*, Ruggero Basanisi[1]*, Matteo Neri[1], Guillaume Auzias[1], Giovanni Petri[2], Daniele Marinazzo[3], Stefano Panzeri[4], Andrea Brovelli[1]

[1] *Institut de Neurosciences de la Timone UMR 7289, Aix Marseille Université, CNRS, 13005, Marseille, France*

[2] *Network Science Institute, Northeastern University, London, London E1W 1LP, United Kingdom*

[3] *Department of Data Analysis, Faculty of Psychological and Educational Sciences, B-9000 Ghent, Belgium*

[4] *Institute for Neural Information Processing, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany*

* co-first author

Corresponding author:

Andrea Brovelli - andrea.brovelli@univ-amu.fr

Keywords

Brain-wise representation, information seeking,

Acknowledgements

A.B. and E.C were supported by the PRC project "CausaL" (ANR-18-CE28-0016) and received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3). A.B. was supported by A*Midex Foundation of Aix-Marseille University project "Hinteract" (AMX-22-RE-AB-071). R.B. and M.N. have received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency (reference : ANR-16-CONV000X / ANR-17-EURE-0029) and from Excellence Initiative of AixMarseille University - A*MIDEX (AMX-19-IET-004). The "Center de Calcul Intensif of the Aix-Marseille University (CCIAM)" is acknowledged for high-performance computing resources. A.B, E.C and D.M were supported by EU's Horizon

2020 Framework Programme for Research and Innovation under the Specific Grant Agreements No. 101147319 (EBRAINS 2.0 Project). G.A. was supported by the ANR SulcalGRIDS Project (ANR-19-CE45-0014). We thank Pedro Mediano and Luca Faes for fruitful suggestions and discussions.

## ABSTRACT

The ability to create beliefs about the consequences of our actions, also known as goal-directed learning, is a key facet of cognition and it provides the basis for rational decision-making. Goal-directed learning arises from a distributed neural circuit including the prefrontal, posterior parietal and temporal cortices. However, the role of cortico-cortical functional interactions remains unclear. To tackle this question, we combined information decomposition theory with human magnetoencephalography (MEG) and studied whether and how learning signals are encoded through neural interactions. Our findings revealed that 'information gain' - or the decrease in uncertainty regarding the causal relationship between an action and its consequence - is represented within a distributed cortical network, incorporating the visual, parietal, lateral prefrontal, and ventromedial/orbital prefrontal cortices. Remarkably, cortico-cortical interactions encoded information gain in a synergistic manner, beyond what individual regions represented alone. Synergistic interactions encoded information gain at the level of pairwise and higher-order relations, such as triplets and quadruplets. Higher-order synergistic interactions were characterised by long-range relations gravitating over the ventromedial and orbitofrontal cortices, which played a receiving role in the broadcasting of information gain over cortical circuits. Overall, the current study provides evidence that information gain is encoded in both synergistic and higher-order functional interactions, as well as through the broadcasting of information gain signals toward the prefrontal reward circuitry. Moreover, our research offers a new perspective on how information relevant to cognition is encoded and broadcasted within distributed cortical networks and brain-wide dynamics.

# INTRODUCTION

A key factor of human agency is the ability to create beliefs about the consequence of our actions. This ability provides the basis for rational decision-making and, in general, allows people to engage in meaningful life and social interactions all along their life (Bandura, 1997). The acquisition of causal action-outcome relations is thought to be supported by the goal-directed system (Tolman, 1948; Dickinson, 1994; Dolan and Dayan, 2013). Neurally, goal-directed behaviours emerge from the coordinated activity of neural populations distributed over the associative fronto-striatal circuit (Balleine and O'Doherty, 2010; Dolan and Dayan, 2013; Balleine, 2019) and the limbic "reward" system (Haber and Knutson, 2009). More specifically, human neuroimaging and primate neurophysiology have shown that goal-directed learning recruits large-scale cortical circuits including the lateral and medial prefrontal cortex (lPFC and mPFC), posterior parietal cortex, orbito-frontal cortex (OFC), temporal gurus and visual cortex, temporal parietal junction (Tanaka, Balleine and O'Doherty, 2008; Walton *et al.*, 2010; Liljeholm *et al.*, 2011, 2013; Jocham *et al.*, 2016; Averbeck and Costa, 2017; Bartolo and Averbeck, 2020; Norton and Liljeholm, 2020; Averbeck and O'Doherty, 2021; Morris *et al.*, 2022).

Goal-directed learning is rooted in the balance between reward maximisation and information seeking processes (Cohen, McClure and Yu, 2007; Frank *et al.*, 2009; Badre *et al.*, 2012; Gottlieb *et al.*, 2013; Mehlhorn *et al.*, 2015; Schwartenbeck *et al.*, 2019; Cockburn *et al.*, 2022). Reward maximisation is classically formalised using associative (Rescorla, 1991) and reinforcement learning models (Watkins and Dayan, 1992; Sutton and Barto, 2018). In these models, agents learn to associate actions with outcomes to maximise cumulative rewards. Update of action values is driven by reward prediction error (RPE) signals, which indicate whether an outcome is better or worse than expected. At the neural level, RPE signals are encoded by neurons in the midbrain, ventral striatum and ventromedial prefrontal cortex (vmPFC) (Frank, Seeberger and O'reilly, 2004; O'Doherty *et al.*, 2004; Pessiglione *et al.*, 2006; Schultz, 2006; D'Ardenne *et al.*, 2008; Palminteri *et al.*, 2015; Gueguen *et al.*, 2021). In parallel, information seeking processes aim at reducing uncertainty in the causal relation between action and outcome, and they drive learning by means of signals quantifying the degree of surprise and the extent to which new information contributes to learning (Friston *et al.*, 2015; Schwartenbeck *et al.*, 2019). Information gain (IG) is often formalised in model-based RL algorithms or Bayesian

models (Yu and Dayan, 2005; Gläscher *et al.*, 2010; Friston *et al.*, 2015; Faraji, Preuschoff and Gerstner, 2018; Gottlieb and Oudeyer, 2018; Schwartenbeck *et al.*, 2019; Liakoni *et al.*, 2021; Modirshanechi *et al.*, 2023) and it is often quantified using the Bayesian surprise measures (Itti and Baldi, 2009; Baldi and Itti, 2010), which indicate how much the agent's belief changes given the new observation. Classical associative learning models formalise this process using the concepts of "surprisingness" or "associability" of events (Mackintosh, 1975; Pearce and Hall, 1980; Courville, Daw and Touretzky, 2006). At the neural level, information gain has been mapped to the activity of distributed brain networks including the middle frontal gyrus, the insula and the intraparietal sulcus (Gläscher *et al.*, 2010; Lee, Shimojo and O'Doherty, 2014; Fouragnan, Retzler and Philiastides, 2018; Liakoni *et al.*, 2022; Modirshanechi *et al.*, 2023). Neural correlates of subjective value of information in instrumental settings (i.e., information that can be used to guide future actions and future outcomes) have been observed in a distributed network including the ventral striatum, the vmPFC, the middle and superior frontal gyrus (i.e., the dorsolateral prefrontal cortex dlPFC), and posterior cingulate cortex (Kobayashi and Hsu, 2019; Kobayashi *et al.*, 2021; Kobayashi and Kable, 2024).

Overall, accumulating evidence suggests that the brain encodes in a distributed manner update signals supporting reward maximisation and information gain during goal-directed learning. The role of cortico-cortical functional interactions underlying goal-directed learning signals remains, however, elusive. To address this issue, we used information decomposition techniques to investigate whether cortico-cortical functional interactions, defined as statistical relationships between the activity of different cortical regions (Panzeri *et al.*, 2022), encode update signals for reward maximisation and information gain. In particular, we tested two hypotheses that recently emerged about the nature of functional interactions. The first hypothesis suggests that brain circuits carry information about cognitive processes by means of redundant or synergistic functional interactions (Varley, Sporns, *et al.*, 2023; Luppi *et al.*, 2024). The second hypothesis suggests that cognitive functions emerge from higher-order brain interactions beyond pairwise relations (Martignon *et al.*, 2000; Yu *et al.*, 2011; Shahidi *et al.*, 2019; Chelaru *et al.*, 2021; Panzeri *et al.*, 2022; Varley, Pope, Maria Grazia, *et al.*, 2023).

Here, we used an experimental task design that manipulates learning and induces highly reproducible explorative strategies across participants and sessions. At the neural level, we combined network-level information decomposition techniques and

magnetoencephalography (MEG) to test whether goal-directed learning signals associated with reward prediction-error and information gain are encoded in redundant and synergistic cortico-cortical interactions, and whether they engage in higher-order collective behaviours broadcasting information across brain regions. Brain functional interactions were studied by analysing high-gamma activity (HGA, from 60 to 120Hz) using a novel approach based on recent advances in information theory, the so-called partial information decomposition (Williams and Beer, 2010; Wibral *et al.*, 2017; Lizier *et al.*, 2018).

We found that information gain is encoded during goal-directed learning in a distributed and cortical network including the visual, parietal, lateral prefrontal and ventro-medial and orbital prefrontal cortices. Interestingly, information gain was also encoded in redundant and synergistic cortico-cortical functional interactions. These functional interactions displayed complementary local-*versus*-distributed properties, synergistic interactions begin more prone for long-range interactions. Most importantly, we found that information-related synergy displayed higher-order interactions, beyond pairwise relations, with a key role played by the prefrontal reward circuit, which acted as a receiver in the network. Overall, we suggest that higher-order synergistic functional interaction is important for encoding and propagating learning signals related to information seeking during goal-directed learning, with a pivotal role played by the limbic cortical regions.

# RESULTS

### *Behavioural scores, exploration strategy and learning signals*

We first examined learning patterns and behavioural strategies utilised by the participants. Specifically, the experimental setup was structured to methodically regulate the exploratory phase and guarantee consistent performance throughout sessions and among different individuals (Brovelli et al., 2008). Accordingly, the correct stimulus-response associations were not set *a priori*. Instead, they were assigned as learning proceeds in the task (Fig. 1b). The first presentation of each stimulus is always followed by an incorrect outcome, irrespective of subject's choice (Fig. 1b, trials 1 to 3). On the second presentation of stimulus S1, any new untried finger movement was considered as a correct response. For the second stimulus S2, the response was defined as correct only when the subject had performed 3 incorrect finger movements (trial 9 in Fig. 1b). For stimulus S3, the subject had to try 4 different finger movements before the correct response was found (trial 14 in Fig. 1b). In other words, the correct response was the second finger movement (different from the first tried response) for stimulus S1, the fourth finger movement for stimulus S2, the fifth for stimulus S3. This task design assured a minimum number of incorrect trials during acquisition (1 for S1, 3 for S2 and 4 for S4) and fixed representative steps during learning. In particular, it produced reproducible behavioural performances across sessions and subjects and allows in depth analysis of group-level behavioural and fMRI data (Brovelli *et al.*, 2008).

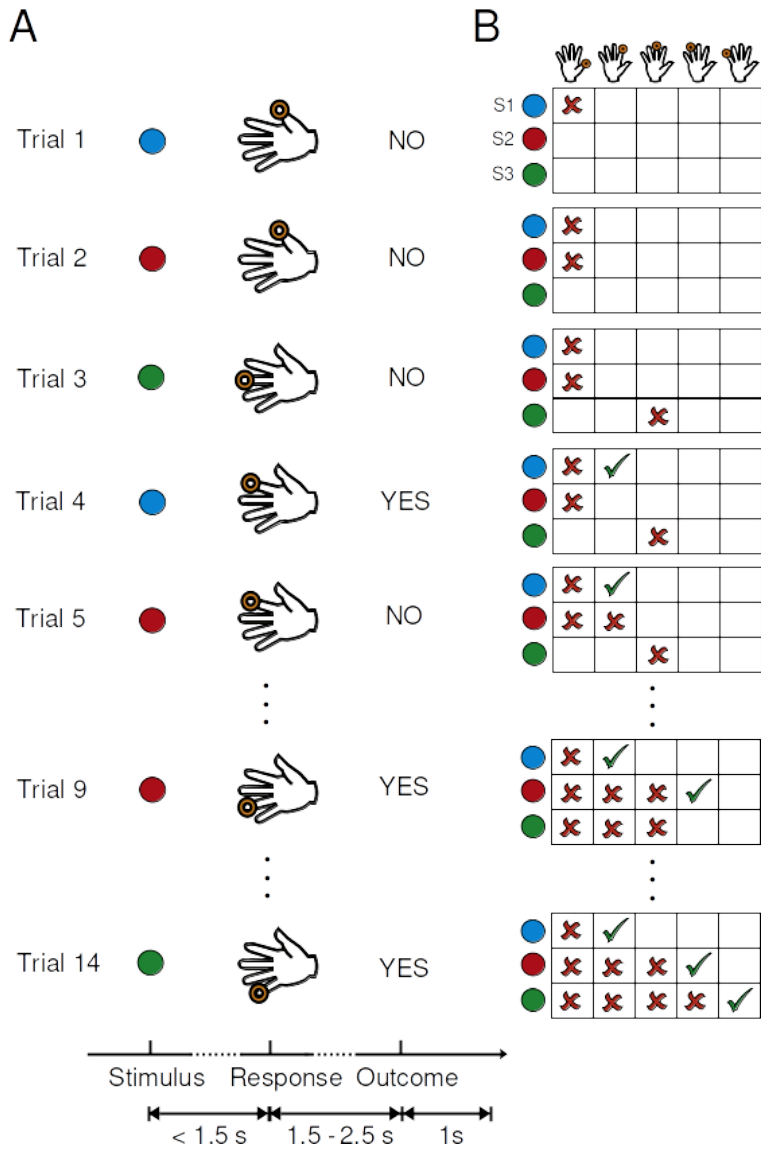


**Figure 1. Experimental design. (A)** On each trial, subjects were presented a coloured circle to which they had to respond within 1.5 s. The outcome image was presented after a variable delay ranging from 1.5 to 2.5 s (randomly drawn from a log-normal distribution). The outcome image informed the subject whether the response was correct, incorrect, or too late (if the reaction time exceeded 1.5 s). **(B)** Matrix of all the possible stimulus–response combinations, updated according to the exemplar learning session in Fig. 1A. A red cross and a green tick-mark refer to incorrect and correct stimulus–response sequences, respectively. The first presentation of each stimulus was always followed by an incorrect outcome, irrespective of the motor response (from trials 1 to 3). On the second presentation of the stimulus S1 (the blue circle), any untried finger movement was always followed by a correct outcome (trial 4). The correct response for S2 and S3 (red and green circles, respectively) was found after 3 and 4 incorrect finger movements (at trials 9 and 14, respectively).

Each participant performed four sessions of 60 trials each (total of 240 trials). Each session included three colours randomly presented in blocks of three trials. The average reaction time was 0.504s ± 0.04s (mean ± s.e.m.). The total number of error trials prior to the first correct outcome was 9.775 ± 0.38 (mean ± s.e.m.), thus indicating a low variability

across participants and sessions. The total number of errors after the first correct outcome was 4.15 ± 0.34, and they represented approximately 9% of the trials during the exploitative phase. This demonstrated that the number of maintenance errors was low.

For what concerns the behavioural strategies, participants employed a standard "tree-search" heuristic, characterised by a tendency to repeat the same finger movement across trials until the correct response was found. This strategy is referred to as "lose-stay", where agents maintain the response when not rewarded. Given that stimuli were pseudo-randomised in blocks of three trials, we computed the probability of repeating a given choice within a block of trials. This probability quantified the likelihood of a directed from random exploration. Indeed, for the first block (trials 1 to 3), the adoption of a lose-stay strategy occurred in 67.5% ± 6%, and in the second block of three trials (trials 4 to 6) occurred in 37.5% ± 4.5%. If the participants had employed a random exploration strategy, the proportion of consecutive trials with identical actions would be equivalent to the ratio of the total number of three-trial patterns with matching actions (n = 5) to the number of possible ordered triplets derived from all combinations of 3 distinct actions chosen from 5 available options (m = 60 combinations). Therefore, the likelihood of encountering a sequence of trials with the same action due to chance alone would be 8.3%. Our results demonstrate that participants used a directed exploration strategy during learning and rapidly attained learning and proficiency during each learning session.

A Q-learning model (Watkins & Dayan, 1992; Sutton & Barto, 1998) was fitted to the behavioural data to estimate changes in stimulus-action-outcome probabilities and learning signals over trials. Reward prediction errors (RPE) and information gain (IG) signals were extracted from the model. While RPE represents the discrepancy between the received and the anticipated outcome, IG is calculated using the Kullback-Leibler divergence measuring the distance between the probability distributions of actions before and after the outcome. RPE embodies a scalar learning signal, fluctuating positive or negative based on whether reality surpasses or falls short of expectations, respectively. Meanwhile, IG captures the magnitude of information conveyed by an outcome, and the amount of update in the associative model.

### *Brain areas encoding information gain signals during learning*

To characterise both the spatial organisation and temporal dynamics of the brain areas encoding learning signals, we performed time-resolved mutual information analyses

between the across-trials high-gamma activity (HGA) and outcome-related learning signals, such as reward prediction errors (RPE) and information gain (IG). Group-level random-effect analyses, combining permutation tests and cluster-based statistics (see Methods), were used to identify temporal clusters displaying significant encoding of learning signals and HGA. Contrary to expectation, we did not find any cortical area significantly encoding ($p_{cluster} < 0.05$) RPEs. On the other hand, a large-scale network comprising visual, temporal, parietal and frontal areas displayed significant encoding of IG (Fig. 2). Forty-two brain regions were found to encode significantly IG, approximately 51% of the brain regions. The strongest and earliest response was observed in the bilateral occipital regions, with a peak centred on the primary visual areas (VCcm and CMrm) of *MarsAtlas* (see top meddle inset of Fig. 2), corresponding to Brodmann areas (BA) 17, 18, 19 and 39. Significant encoding was also observed in the left medial and superior parietal cortices (SPCm and PCm), corresponding to BA 7 and 31, respectively. In the temporal lobe, the right temporal cortices displayed significant encoding of IG. In the right frontal lobe, the right premotor areas (PMdl and PMdm), BA 6 and 8, and the right dorsolateral prefrontal cortices (PFcdm, PFcdl, Pfrdli and Pfrdls), PFdls, PFrdl and PFrdll), roughly corresponding to BA 45, 46, 27 and the most posterior parts of BA 9 and 10. Finally, the bilateral activation of the ventromedial prefrontal cortex (PFCvm, BA 32/10/11), orbitofrontal areas (OFCv, OFCvl and OFCvl, roughly BA 10/11/47) and the insula were found to display significant encoding of IG. Broadly speaking, IG was encoded in a set of height cortical subnetworks (or clusters) including bilateral visual areas, the right temporal regions, the left superior parietal areas, the right motor and premotor areas, the right dorsolateral prefrontal cortices and the bilateral ventromedial prefrontal and orbitofrontal areas.

In order to characterise more specifically the time course of learning-related signals encoding IG, we averaged the mutual information time courses across the height clusters of cortical regions displaying significant effects (Fig. 3). The average time courses of learning-related modulations for the height cortical clusters are shown in Fig. 3. The visual (Fig. 3A) and parietal (Fig. 3B red curve) clusters displayed a similar time course characterised by a fast and transient increase in MI and peaking around 0.2 after outcome onset.

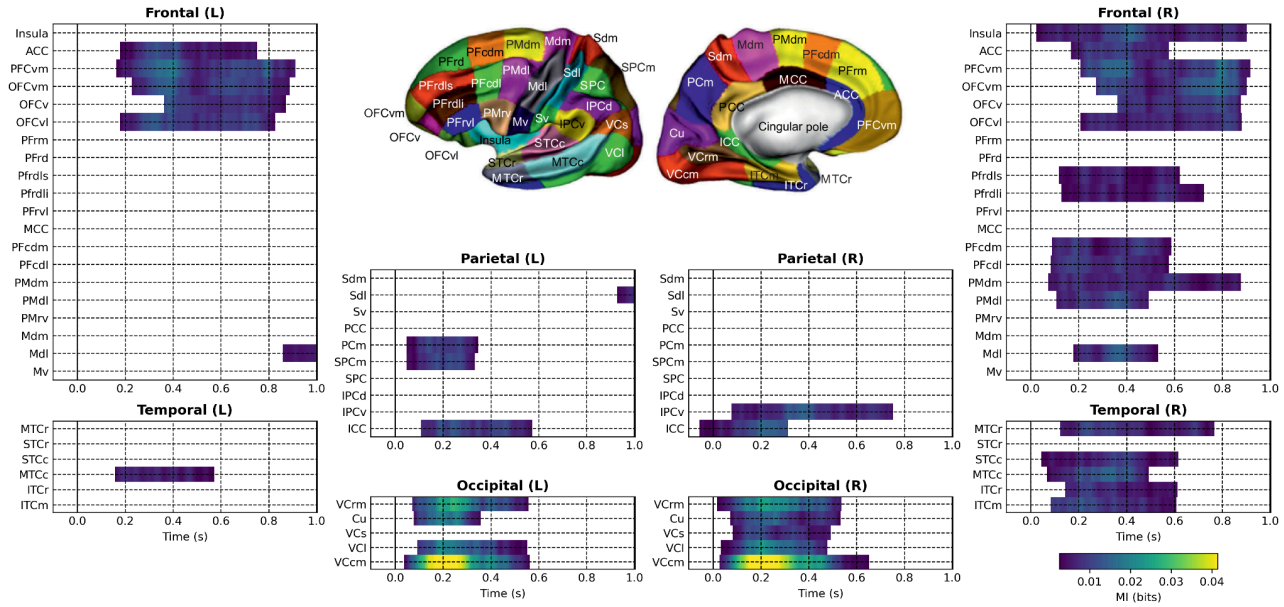


**Figure 2.** Local spatio-temporal correlates of information gain. Regions-of-interest (ROI) of the MarsAtlas (Auzias et al., 2017) are shown in the inset in the top-central part of the figure. ROIs are grouped for the different lobes (occipital, parietal, temporal and frontal) and hemisphere (left and right panels). The panels in the figure show the temporal cluster displaying significant encoding of IG(p_cluster < 0.05). Colormap shows the value of the mutual information (in bits) between the HGA and IG.

The encoding in temporal regions appeared later in time than in visual and parietal regions, peaked around 0.3s after outcome onset (Fig. 3B, blue curve) and returned to based around 0.8s after outcome onset. The premotor and dorsolateral regions of the right hemisphere peaked approximately from 0.3 to 0.4s after outcome and they were characterised by a slower return to baseline (Fig., 3C). Finally, the vmPFC and OFC regions were characterised by a late onset at approximately 0.2s and they displayed a sustained encoding lasting until 1s (Fig. 3D).

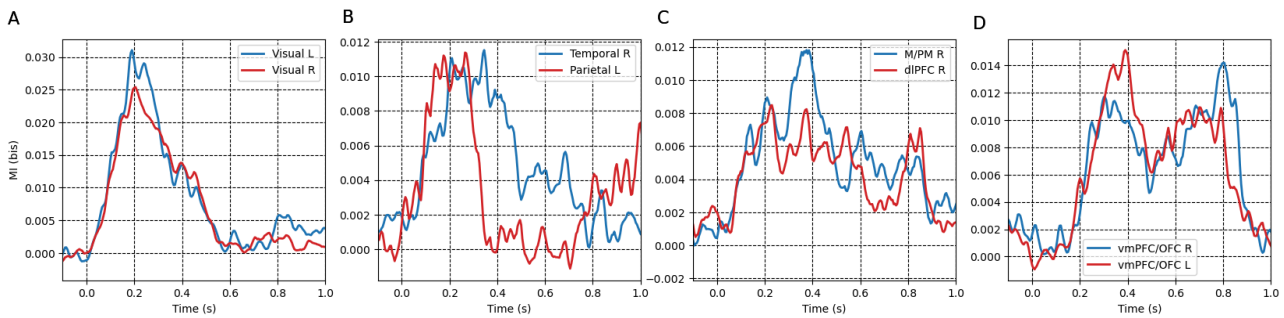


**Figure 3.** Time courses of mutual information values for the eight clusters of cortical areas significantly encoding information gain. A) Left (blue) and right (red) visual areas; B) temporal (blue) and parietal (red) areas; C) premotor (blue) and dorsolateral prefrontal (red) areas; D) left (blue) and right (red) ventromedial prefrontal and orbitofrontal areas.

To differentiate complementary spatio-temporal patterns of encoding of IG, we performed non-negative matrix factorisation (NMF) of the MI values across clusters. Figure 4A shows the NFM using 4 components, defined using the so-called the elbow method, which is a standard heuristic for determining the number of components in a dataset. Figure 4B shows the variance explained by each component (in colour) for each brain area (rows). The first component (in blue) modelled the fastest and most transient encoding of IGand it is mainly observed in the visual and parietal areas (blue bars in Fig., 4B). The orange and green components occurred later in the trial peaking around 0.35 and 0.45s after outcome onset and they recruited primarily the temporal, premotor and dorsolateral areas. Finally, the fourth component (in red) captured the slowest component that emerged around 0.5s and later approximated 0.4s. This latest component mainly recruited the ventromedial PFC and OFC cluster. Overall, these results show that the encoding of IGrecruits a large-scale network of cortical clusters with overlapping temporal and spatial dynamics.

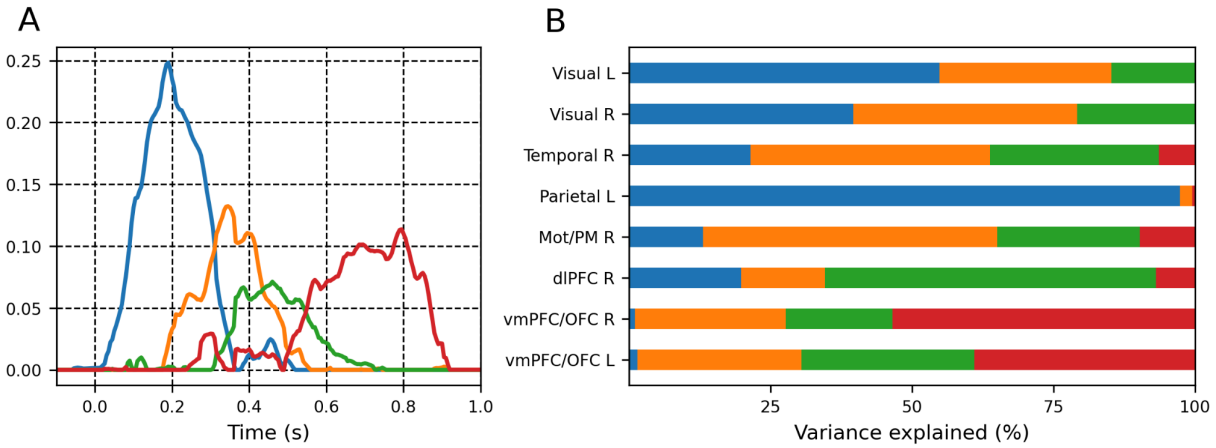


**Figure 4.** (A) Non-negative matrix factorisation loading. (B) Explained variance associated with each loading for all brain regions.

## *Cortico-cortical interactions mediating information gain*

We then tested the hypothesis that learning computations such as information gain emerge from network-level and cortico-cortical interactions in addition to local activations. To do so, we exploited recent advances in information theory that allow isolating the contribution of network-level interactions from those carried individually. Most importantly, we aimed at characterising networks associated with functional segregation and integration processes by decomposing the information carried by pairs of brain regions

about IG into redundant (i.e., shared or common between regions) and synergistic (i.e., requiring the joint combination), respectively. Significant redundant cortico-cortical co-modulations encoding IG were observed over a large-scale network (Fig., 5A). Strongest redundant connectivity was observed earlier over bilateral visual areas during the first 0.2s after outcome presentation (left panel of Fig., 5A). Redundancy connectivity then spread over the right temporal lobe, right motor and premotor cortices, and ventromedial PFC and OFC cluster in the time interval between approximately 0.2 and 0.8s. Redundant connectivity over the ventromedial PFC and OFC cluster were long lasting and faded away approximately 0.8 to 10s after outcome (right panel of Fig., 5A). In addition, we observed significant synergistic interactions between pairs of brain regions (Fig. 5B). Differently from redundant connectivity, synergistic connectivity appeared to link distant areas of the visual system, the right temporal areas, the right dorsolateral PFC and the left ventromedial PFC and OFC. The peak of interaction strength occurred from 0.2 to 0.5s after outcome.

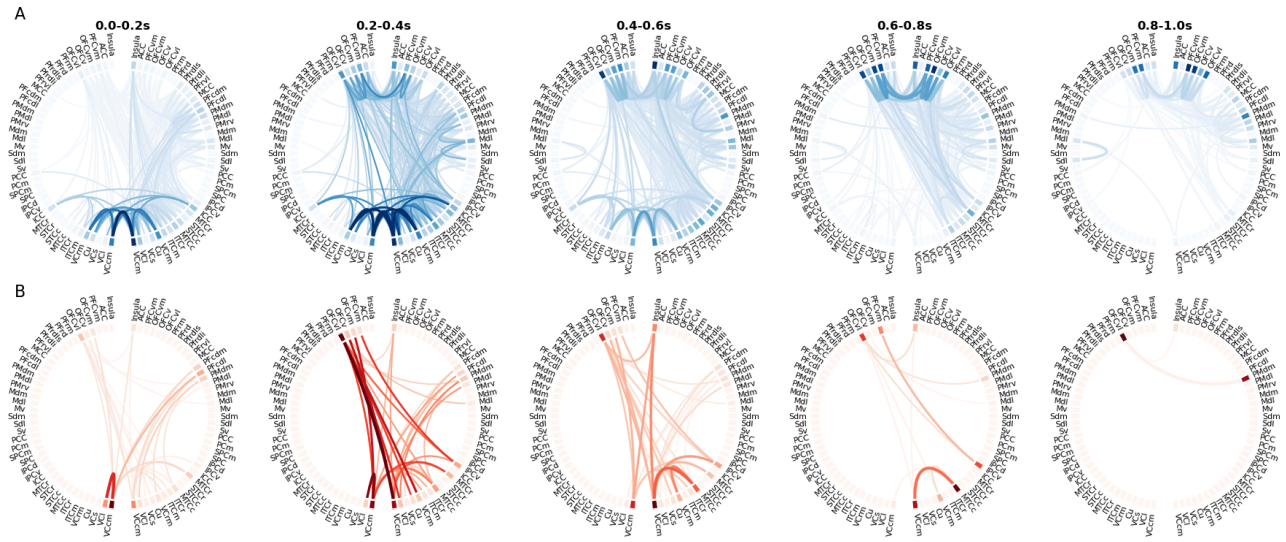


**Figure 5.** Redundant (A) and synergistic (B) pairwise interactions significantly carrying information about information gain signals. Each panel is the average value in 0.2s time windows aligned on outcome presentation. The labels of the regions of interest plotted in a circle are identical to those of Fig. 2. The position of each brain regions over the cortical sheet and lobe can be seen in the inset of Fig. 2

From visual inspection, redundant connectivity appeared to predominantly recruit pairs of brain regions with the same system (or cluster), whereas synergistic connectivity appeared to be stronger across systems. In order to quantify such an effect, we computed the average redundant and synergistic connectivity between pairs of brain regions within or across eight clusters of cortical areas significantly encoding information gain, respectively. Indeed, we found that redundant connectivity showed stronger average value

across pairs of brain regions within the same cluster (Fig., 6A), whereas synergistic coupling showed a stronger effect across clusters (Fig., 6B).

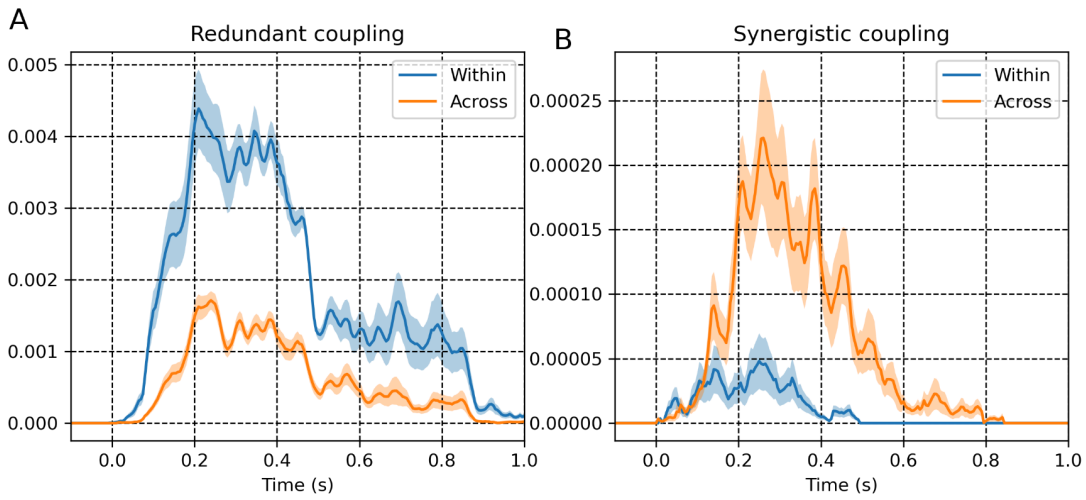


**Figure 6.** Average redundant (A) and synergistic (B) connectivity computed over pairs of brain regions being part of the same cluster or different cluster. The "within-cluster" connectivity is plotted in blue, whereas the "across-cluster" connectivity is in orange. Mean and standard error of the mean are plotted as solid line and shade area, respectively.

By definition (Eq. 9), redundant functional interactions carry the minimum information about learning variables carried by individual brain areas. On the other hand, synergistic connectivity reveals brain functional interactions that cannot be explained by individual brain areas, but emerge from the collective coordination between areas. Significant synergistic functional interactions therefore reveal the presence of collective behaviour encoding IG tha emerge from the coordination of distant brain areas and circuits.

### *Higher-order cortico-cortical interactions and information gain*

The previous result showed that synergistic interactions play a role across visual, prefrontal and temporal areas. This suggests that large-scale functional integration processes encoding information gain emerge between distant cortical clusters through synergistic interactions. This highlights the important role of collective and synergistic processes beyond the contribution of individual brain areas and redundant information encoding. Since we focused on regions of interest locally encoding information gain, the presence of higher-order redundancy trivially arises from lower order (e.g., pairwise) relations. This is due to the definition of redundancy, which is equal to the minimum of the local mutual information values computed between the brain activity and information gain. Thus, we exclusively investigated the role of synergistic coupling in the search for neural

markers of functional integration phenomena beyond pairwise relations. Inspired by recent studies evidencing higher-order correlations beyond pairwise relations as a potential substrate for emergent phenomena (Gatica et al., 2021; Herzog et al., 2022; Varley, Pope, Maria Grazia, et al., 2023), we investigated whether cortico-cortical interactions encode information gain through higher-order co-modulations. To do so, we studied higher-order relations between nine cortical clusters encoding IG. The clusters were associated with the left and right visual areas (VCcm), the left parietal (IPCv), the right temporal cortices (MTCc and MTCr), the right motor area (Mdl), right premotor areas (PMdl and PMdm), right dorsolateral prefrontal areas (PFcdm and PFcdl) and the left and right ventromedial and orbitofrontal areas (OFCvl, OFCv, OFCvm and PFCvm), as shown in Fig. 2. Given that the clusters encoded individually information gain, it is trivial that they also encode redundant information at the higher-order relation level. However, higher-order synergistic connectivity, which represents information emerging from collective patterns, warrants further investigation. Cluster-based statistical analyses and correction for multiple comparisons (across all pairs and multiples across all orders) were performed to identify significant higher-order encoding of IG. A total of 19 pairs, 12 triplets and 2 quadruplets were found to significantly encode IG (Fig. 7A). Higher-order synergistic co-modulations carrying information about IG were found up to order 4. Given that we analysed 9 clusters, the percentage of significant links, triplets and quadruplets was 53%, 14% and 2%, respectively. We next asked whether higher-order functional interactions beyond pairwise relations were associated with hypergraph or simplicial complex representations. As a reminder, a hypergraph is a description of higher-order interactions, consisting of k nodes representing a k-way interaction between the nodes. A simplicial complex is a special case of a hypergraph, where all possible (k-1)-way interactions among that same set of nodes occur (Battiston *et al.*, 2020, 2021). In other words, hypergraphs without lower-level interactions may reflect a stronger “marker” of emergent collective behaviour, because it would lack lower level components. Although the observed higher-order interactions pertain to the functional rather than structural case, we investigated whether higher-order functional interactions are better described by hypergraphs or simplicial complexes. We found that 7 out of 12 triplets were composed of significant pairwise links, and therefore reflected simplicial complexes (Fig. 7A). Five triplets were, however, composed of nodes that did not display all pairwise interactions, therefore reflecting a stronger emergent form of collective behaviour. These non-simplician triplets all included the visual areas (either

left or right VCcm) and the left ventromedial and orbitofrontal areas (left mvPFCOFC), which were associated with significant pairwise relations. The “missing” pairwise links occurred between the left ventromedial-orbitofrontal areas, and the left parietal, right premotor and right dorsolateral areas, therefore highlighting that such triplets encoded IG at order 3. Similarly, the two quadruplets significantly encoding IG were not composed of all lower order triplets, and they involved the visual-temporal-vmPFCOFC and the visual-parietal-vmPFCOFC clusters. Overall, these results demonstrate that IG is encoded in a distributed manner across synergistic interactions beyond local brain regions and shows genuine higher-order synergistic interactions.

Finally, we characterised the temporal evolution of such higher-order synergistic interactions. Significant temporal clusters were mainly observed over the interval between 0.1s and 0.6s, and no clear differentiation between pairwise and 3rd- or 4th-order interactions. In order to better characterise the participation of clusters in each multiple, we computed the synergistic weight defined as the summed synergistic connectivity associated with a given cluster and we studied how it evolved over time. Figure 7B shows the synergistic weight for pairwise relations, and it highlights that the bilateral visual cluster dominates the interaction pattern. However, for higher orders the visual clusters were accompanied by the left vmPFC-OFC dominated both for triplets (Fig. 7C) and quadruplets (Fig. 7D).

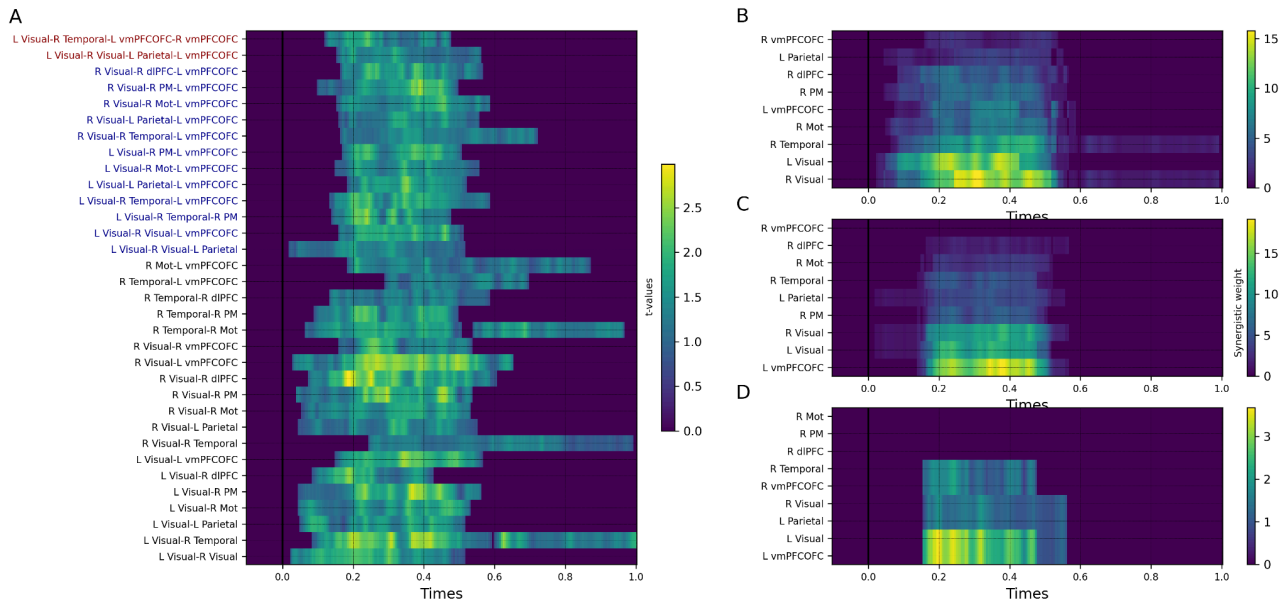


**Figure 7.** Synergistic higher-order interactions between cortical clusters encoding information gain (IG). (A) Time course of t-values showing significant encoding of IG at the level of pairs of clusters (black labels),

riplets (blue label) and quadruplets (red labels). (B) Cluster weight for pairwise interactions. (C) and (D) show cluster weight for triplets and quadruplets, respectively.

### *Cortico-cortical information transfer encoding learning signals*

Previous results highlighted a time-lagged encoding of IG from visual to prefrontal and vmPFC areas, suggesting a broadcasting of IG signals to the vmPFC. In order to test this hypothesis, we exploited a recently-developed information-theoretic measure termed Feature-specific Information Transfer (FIT). The FIT measure merges the Wiener-Granger causality principle with content specificity and it allows tracking task-related directional includes between neural signals (Celotto et al., 2023). We used the FIT measure to quantify how much information about IG between the nine clusters of brain regions displaying higher-order synergistic encoding. Four directional relations significantly encoded IG and broadcasted information to the ventromedial PFC and OFC (Fig. 8). The temporal information is aligned from the point of view of the receiving brain area. The IG-specific information transfer peaked around 0.35 to 0.4s after outcome onset, thus roughly corresponding to the first peak in HGA observed locally. Two patterns of convergent interactions towards the left and right vmPFCOFC, respectively. These “colliding” patterns included two higher-order synergistic multiplet, namely a triplet including the right visual and temporal regions with the left vmPFCOFC, and a quadruplet including ight visual and temporal regions with the bilateral vmPFCOFC (Fig. 7).

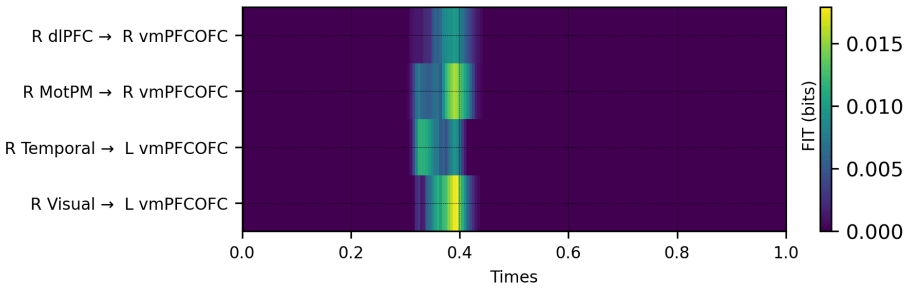


**Figure 8.** Feature-specific Information Transfer (FIT) encoding Information Gain (IG). Significant directional interactions broadcast information about IG primarily to the vmPFCOFC.

# Discussion

## Information seeking and directed exploration in goal-directed learning

Goal-directed learning is defined as the ability to learn the causal relationship between our behaviours and their outcomes, and to select actions according to expected outcomes, as well as current goals and motivational state (Tolman, 1948; Dickinson, 1994; Dolan and Dayan, 2013). Goal-directed learning relies on the balance between exploitation and exploration strategies (Cohen, McClure and Yu, 2007; Frank *et al.*, 2009; Badre *et al.*, 2012; Gottlieb *et al.*, 2013; Schwartenbeck *et al.*, 2019; Cockburn *et al.*, 2022). A trade-off between choosing the most valuable option and sampling informative options is central for successful goal-directed learning and it is one of the main focus of reinforcement learning and Bayesian learning theories (Friston *et al.*, 2015, 2017; Sutton and Barto, 2018). Exploration in humans has been shown to be based on a mixture of random and directed strategies (Wilson *et al.*, 2014; Gershman, 2018). Random exploration is a deviation from the most rewarding option, and it is normally formalised using the softmax choice rule, which models the degree of randomness (decision noise). Directed exploration selectively samples options that are informative (information seeking), that are associated with the highest uncertainty.

Here, we have addressed the neural computations supporting information seeking processes during goal-directed learning in terms of information gain and directed exploration. In order to investigate directed exploration and information gain, we adapted a learning task previously developed for fMRI (Brovelli et al., 2008) that manipulates learning and induces reproducible explorative strategies across subjects and sessions. Indeed, we observed that learning was characterised by a directed exploration strategy based on a “tree-search” or “lose-stay” pattern, in which participants had a tendency to repeat the same action after incorrect outcomes and to span all options in an ordered manner. Given that the correct action-outcome association was deterministic, such directed exploration strategy pertained only the initial phase of the task, up to the first correct outcome. Subsequently, participants performed a relatively low number of maintenant errors and engaged in an exploitative strategy. The presence of lose-stay strategy demonstrates the directed nature of the explorative behaviour during the learning task. From the computational point of view, information seeking and exploration is formalised using model-based RL algorithms or Bayesian models (Yu and Dayan, 2005; Gläscher *et al.*,

2010; Friston *et al.*, 2015, 2017; Faraji, Preuschoff and Gerstner, 2018; Liakoni *et al.*, 2021; Modirshanechi, Brea and Gerstner, 2022) and it has been proposed to play a more general role in curiosity-driven learning (Gottlieb and Oudeyer, 2018; Schwartenbeck *et al.*, 2019). Therefore, the total subjective value of actions does not exclusively depend on the expected subjective utility (or extrinsic value), but is also depends on the informative (intrinsic, epistemic or non-instrumental) value expected to be gained and which leads to a reduction in uncertainty about environmental contingencies (Cohen, McClure and Yu, 2007; Bromberg-Martin and Hikosaka, 2009; Tishby and Polani, 2011; Friston *et al.*, 2015; White *et al.*, 2019). In other words, choices are thought to be made through the comparison between action "values"; action values can be decomposed in (at least) two terms, which quantify the extrinsic and epistemic value, respectively. For what concerns the update of action values through the processing of outcomes during learning, we suggest that directed exploration is associated with at least two processes: i) the update of action expected reward (or extrinsic) values by means of reward prediction errors (RPE); ii) the update of (intrinsic or informative) action values by means of an information gain (IG) signal. IG signals quantify the trial-by-trial evolution of information gain during directed exploration. Such information is used to decrease the uncertainty in the causal relation between action and outcomes, and reduce the agent's uncertainty about the world. Within the taxonomy of surprise definitions, our study addresses the subcategory defined as "belief-mismatch surprise" (Modirshanechi, Brea and Gerstner, 2022), referred here simply as information gain. A limitation of our study may therefore be the inability to associate the neural results to alternative metrics of information gain within the category of belief-mismatch surprise measures, such as the minimised free energy (Schwartenbeck *et al.*, 2013).

## Distributed spatio-temporal cortical activations encoding information gain

We thus investigated how the brain updates the subjective value of choices based on information gain. To study the neural substrates of information gain, we performed information theoretical analyses of single-trial and cortical-level high-gamma activity (HGA) from MEG. A large-scale network comprising visual, temporal, parietal and frontal areas encoded information gain (IG) during goal-directed learning (Fig. 2). We will now discuss the relevance of the current results at different levels from entire brain circuits, to hemispheric and regional contributions.

At the circuit level, the results suggest that information gain is encoded in the dorsal and ventral circuits of the goal-directed system (Averbeck B, 2020; Averbeck and O'Doherty, 2021). The dorsal circuit, comprising the dorso-lateral prefrontal cortex (dlPFC) and the inferior parietal cortices, is thought to primarily encode information relevant for actions and control behaviour to achieve goals. The ventral circuit, comprising the ventro-medial prefrontal cortex (vmPFC) and orbito-frontal cortex (OFC), is suggested to primarily learn to assign and update subjective values to potential outcomes and motivational value of objects in the environment, owing to visual input from the anterior temporal cortex and connections with the hypothalamus (Averbeck and Costa, 2017). Our results show that information gain associated with outcomes is encoded over both the ventral and dorsal circuits. Therefore, we suggest that previous human neuroimaging and primate neurophysiology studies highlighting learning-related effects over the lateral and medial prefrontal cortex (lPFC and mPFC), posterior parietal cortex, orbito-frontal cortex (OFC), temporal gurus and visual cortex, temporal parietal junction (Tanaka, Balleine and O'Doherty, 2008; Walton *et al.*, 2010; Liljeholm *et al.*, 2011, 2013; Jocham *et al.*, 2016; Averbeck and Costa, 2017; Bartolo and Averbeck, 2020; Norton and Liljeholm, 2020; Averbeck and O'Doherty, 2021; Morris *et al.*, 2022) may be linked to information gain processes. More specifically, our findings confirm that information gain and Bayesian surprise signals are represented in distributed brain networks including the middle frontal gyrus, the insula and the intraparietal sulcus (Gläscher *et al.*, 2010; Lee, Shimojo and O'Doherty, 2014; Fouragnan, Retzler and Philiastides, 2018; Liakoni *et al.*, 2022; Modirshanechi *et al.*, 2023). Indeed, fMRI activity encoding belief updates (Bayesian surprise) has been observed in the midbrain and ventral striatum, in addition to the presupplementary motor area (pre-SMA), dorsal anterior cingulate cortex, posterior parietal cortex and lateral prefrontal cortex (O'Reilly *et al.*, 2013; Nour *et al.*, 2018). Furthermore, our results complement fMRI studies showing correlates of information value in instrumental settings (i.e., information that can be used to guide future actions and future outcomes) in the ventral striatum, vmPFC, the middle and superior frontal gyrus (i.e., the dorsolateral prefrontal cortex dlPFC), and posterior cingulate cortex (Kobayashi and Hsu, 2019; Kobayashi *et al.*, 2021; Kobayashi and Kable, 2024). Further work is needed to fully appreciate the relation between the processing of information about outcomes and action selection during information seeking.

At the hemispheric level, a right-lateralized activation was observed (Fig. 2). This result

may relate to established literature showing a right-hemisphere dominance for stimulus-driven shifts of spatial attention and target detection (Shulman *et al.*, 2010; Corbetta and Shulman, 2011) and activation patterns observed in the ventral attentional network (Vossel, Geng and Fink, 2014). This result also resonates with associative learning models formalising attentional processes during learning using concepts such as "surprisingness" or "associability" of events (Mackintosh 1975; Pearce and Hall 1980; Courville et al. 2006). We thus suggest that attentional processes and the associated neural circuits play a role in information seeking, whose relation remains elusive.

At the regional level, a key finding was the encoding of information gain in vmPFC and OFC. This suggests a strong role of the limbic cortical regions classically associated with the encoding of reward prediction error (RPE) signals (Frank, Seeberger and O'reilly, 2004; O'Doherty *et al.*, 2004; Pessiglione *et al.*, 2006; Schultz, 2006; D'Ardenne *et al.*, 2008; Palminteri *et al.*, 2015; Gueguen *et al.*, 2021). Our results are in line with the notion of a "common currency" for information and reward values in the limbic system (Levy and Glimcher, 2012; Kobayashi and Kable, 2024). Indeed, these areas are main targets of dopaminergic projections, which encode, in non-instrumental settings, expected advance information about future outcomes encoding errors in information prediction (information prediction error, reflecting the difference between obtained and expected information gain) (Bromberg-Martin and Hikosaka, 2009, 2011; Bromberg-Martin *et al.*, 2024), in the OFC in human fMRI (Charpentier, Bromberg-Martin and Sharot, 2018) and in frontal EEG electrodes (Brydevall *et al.*, 2018). The anterior cingulate cortex (ACC) and two subregions of the basal ganglia (the internal-capsule-bordering portion of the dorsal striatum and the anterior pallidum) signalling reward uncertainty and information-anticipatory activity (White *et al.*, 2019) may additionally provide a substrate for common encoding of reward and information value. The underlying neural mechanisms are currently unknown, However, converging evidence supports that subpopulations of midbrain dopamine neurons encoding motivational values (Bromberg-Martin, Matsumoto and Hikosaka, 2010) may contribute to the observed activations in limbic regions. This may explain the assignment of value to uncertainty reduction, integrating the value of information with the value of physical rewards, and may provide a link with the neural circuits for information seeking and curiosity-driven behaviours (Monosov, 2024).

More generally, we suggest that the current results provide a deeper understanding of the role of ventromedial PFC and OFC in decision-making and information seeking. Indeed,

it is widely accepted that vmPFC/OFC plays a key role in learning and decision making by encoding outcome information and value (Padoa-Schioppa and Assad, 2006; Ballesta *et al.*, 2020; Gardner *et al.*, 2020; Gore *et al.*, 2023). Moreover, during goal-directed exploration, fMRI activity in the rostral frontopolar cortex (RFPC) activity (Badre and D'Esposito, 2009) and frontal theta component in EEG have indicated a direct role in directed exploration (Cavanagh *et al.*, 2012; Cavanagh and Frank, 2014), rather than random exploration (Daw, Niv and Dayan, 2005). Stimulation and inhibition of RFPC with direct current (tDCS) has shown to increase and decrease the frequency of exploratory choices (Raja Beharelle *et al.*, 2015; Zajkowski, Kossut and Wilson, 2017). These studies put forward the idea that the RFPC is crucial for integrating information about past, present, and future payoffs to arbitrate between exploration and exploitation in human decision making. Our results suggest that a key computation supporting directed exploration and information seeking is the processing of information gain. Altogether, our study strongly supports recent computational models formalising both information gain and reward maximisation processes for learning and decision-making (Friston *et al.*, 2015; Schwartenbeck *et al.*, 2019). Nevertheless, it remains unclear how reward and information maximisation during learning are implemented in neural circuits and network interactions.

Our results provide novel insight into the temporal dynamics information gain during directed exploration, which was lacking in previous neuroimaging and neurophysiological studies in primates. The strongest and earliest encoding of information gain occurred in occipital regions, followed by the parietal and temporal cortices, and the right frontal lobe including premotor and dorsolateral prefrontal cortices, and final activation of the ventromedial prefrontal cortex orbitofrontal areas and insula (Fig. 3). We statistically dissociated four complementary spatio-temporal patterns of HGA encoding IG (Fig. 4). The first component modelled the fastest and most transient encoding of IG recruiting the visual and parietal areas. The second and third component occurred later in the trial peaking around 0.35 and 0.45s after outcome onset and they recruited primarily the temporal, premotor and dorsolateral areas. Finally, the fourth component captured the slowest component that emerged around 0.5s and later approximated 0.4s, involving the ventromedial PFC and OFC cluster. Overall, our results demonstrated that the encoding of IG recruits a large-scale network of cortical clusters with overlapping temporal and spatial dynamics, spanning visual to limbic frontal areas.

## Redundant and synergistic functional interactions encode information gain

We next investigated whether and how information gain is encoded in cortico-cortical interactions. A central hypothesis in neuroscience posits that cognitive functions arise from the coordinated activity of neural populations distributed over large-scale brain networks (Varela *et al.*, 2001; Bressler and Menon, 2010) and from the emergence of network-wide and self-organised information routing patterns (Deco *et al.*, 2015; Battaglia and Brovelli, 2020). Increasing evidence from large-scale measurements of neural activity in mice is confirming the brain-wide nature of task-relevant representations (Allen *et al.*, 2019; Steinmetz *et al.*, 2019; Stringer *et al.*, 2019; International Brain Laboratory *et al.*, 2023; Chen *et al.*, 2024). Learning is no exception, and it is thought to reflect a network-level phenomenon emerging from neural interactions distributed over cortical-subcortical circuits (Averbeck and Costa, 2017; Bassett and Mattar, 2017; Hunt and Hayden, 2017; Averbeck and O'Doherty, 2021). In mice, recent evidence suggests that transformations linking sensation to action during learning and decision-making are highly distributed and parallelized in a brain-wide manner (Khilkevich *et al.*, 2024). Indeed, we confirmed that learning computations emerge from network-level and cortico-cortical interactions (Fig. 5), in addition to local activations (Fig. 2). We performed information decomposition analyses to quantify the information carried by pairs of brain regions, by dissociating the encoding of IG into redundant (i.e., shared or common between regions) and synergistic (i.e., requiring the joint combination) interactions. We observed that IG was encoded both by redundant and synergistic cortico-cortical co-modulations (Fig. 5). Strongest redundant connectivity was observed earlier over bilateral visual areas during the first 0.2s after outcome presentation and then spread over the right temporal lobe, right motor and premotor cortices, and ventromedial PFC and OFC clusters. By definition (Eq. 9), redundant functional interactions equate the minimum of information carried locally. Redundant functional connectivity therefore reflects the shared information carried by co-modulations in HGA across areas (Fig. 5A). On the other hand, synergistic connectivity analysis reveals functional interactions that cannot be explained by individual brain areas, but that emerge from the collective coordination between areas at a specific. Synergistic functional interactions therefore reveal the presence of collective behaviour encoding IG that emerge from the coordination of distant brain areas and circuits (Fig. 5B). Synergistic interactions globally reassembled redundant connectivity patterns. However, contrary to redundant connectivity, synergistic coupling showed a stronger effect across clusters,

rather than within clusters (Fig. 6), therefore suggesting a more distributed and long-range nature of synergistic functional interactions.

We suggest that redundant and synergistic functional interactions could be interpreted as a proxy of functional segregation and integration processes, although a direct mapping may be lacking. Redundant functional interactions may appear in collective states dominated by oscillatory synchronisation (Engel, Fries and Singer, 2001; Varela *et al.*, 2001; Buzsáki and Draguhn, 2004; Fries, 2015) or resonance phenomena (Vinck *et al.*, 2023). Synergistic functional interactions may be associated with functionally-complementary interactions (i.e., functional integration). Indeed, synergistic interactions have been reported between distant transmodal regions during high-level cognition (Luppi *et al.*, 2022; Varley, Pope, Faskowitz, *et al.*, 2023) and, at the microscale, in populations of neurons within a cortical column of the visual cortex and across areas of the visuomotor network (Nigam, Pojoga and Dragoi, 2019; Varley, Pope, Maria Grazia, *et al.*, 2023). During learning, a recent study has shown that redundant functional interactions encoding either reward or punishment prediction errors are associated by segregated networks in prefrontal regions, whereas the integration between reward and punishment learning is mediated by synergistic interactions between them (Combrisson *et al.*, 2023). More generally, synergistic interactions may provide functional advantage, in contrast to redundancy and unique information, because they would enable the full exploitation of possible combinations of neurons or brain areas, making their informational capacity to be exponential with the system size (Rosas *et al.*, 2020). Redundancy, instead, would provide robustness, because over-representation would ensure that information remains available even if any neuron or brain region would be perturbed. A drawback of redundancy would be the cost of over-representation, whereas a limitation of synergistic encoding would be the vulnerability to disruptions in individual nodes would disrupt information synergistically held together with other sources (Mediano *et al.*, 2022). Indeed, information decomposition analysis of artificial neural networks performing cognitive tasks has shown that while redundant information is required for robustness to perturbations in the learning process, synergistic information is used to combine information from multiple modalities and more generally for flexible learning (Proca *et al.*, 2024). Along the same reasoning, artificial networks exhibit higher levels of synergy at comparatively lower orders, rather than higher orders, probably as an attempt to minimise the effect of losing synergistic information with several neurons (Tax, Mediano and Shanahan, 2017). In brain networks,

redundant functional interactions may play a role for robust sensory and motor functions, whereas synergistic interaction for flexible higher cognition (Luppi *et al.*, 2024), and may support system robustness and resilience at the large-scale level, as suggested for microscopic neural networks (Panzeri *et al.*, 2022). Within this interpretative framework, our results provide supports to the hypothesis that brain interactions regulate segregation and integration processes to support cognitive functions (Sporns, 2013; Braun *et al.*, 2015; Deco *et al.*, 2015; Cohen and D'Esposito, 2016; Shine *et al.*, 2016; Finc *et al.*, 2020; Wang *et al.*, 2021), balancing trade-offs between redundancy for robust sensory and motor functions and synergistic interaction for flexible higher cognition (Luppi *et al.*, 2024). Such balanced states may additionally give rise to selective patterns of information flow (Buehlmann and Deco, 2010; Kirst, Timme and Battaglia, 2016; Palmigiano *et al.*, 2017; Battaglia and Brovelli, 2020), as observed in the feature-specific information transfer (FIT) results (Figure 8). Taken together, our results suggest general principles for network interactions whereby synergistic relations encode relevant information for goal-directed learning. The causal effect of synergistic interactions on cognition remains, however, an open question that could be tackled with perturbation or pharmacological manipulations.

## Higher-order synergistic functional interactions encoding information gain

In order to further explore the nature of synergistic cortico-cortical interactions in directed exploration, we tested whether neuronal interactions supporting cognitive functions require collective behaviours among multiple brain areas, the so-called higher-order correlations (Martignon *et al.*, 2000; Schneidman *et al.*, 2006; Montani *et al.*, 2009; Ohiorhenuan *et al.*, 2010; Ganmor, Segev and Schneidman, 2011; Yu *et al.*, 2011; Shimazaki *et al.*, 2012; Nigam, Pojoga and Dragoi, 2019; Sizemore *et al.*, 2019; Panzeri *et al.*, 2022). Indeed, higher-order interactions are at the core of complex systems (Crutchfield, 1994; Battiston *et al.*, 2020, 2021; Santoro *et al.*, 2023) and they are thought to support collective behaviours in social systems (Benson, Gleich and Leskovec, 2016), ecology (Grilli *et al.*, 2017; Levine *et al.*, 2017) and biology (Sanchez-Gorostiaga *et al.*, 2019). By means of information theoretical decomposition methods generalised to higher-order functional interactions, we showed that information gain is encoded in higher-order synergistic functional interactions up to quadruplets. By definition (eq. 12), significant multiplets at a given order carry synergistic information that is not encoded at a lower order. That means, for example, that significant triplets carry information in excess

to the one carried by the underlying pairwise links. The same holds for the two largest quadruplets involving the visual-temporal-vmPFC-OFC and the visual-parietal-vmPFC-OFC clusters (Fig. 7A). Our results therefore demonstrate that IG is encoded by functional interactions beyond the pairwise regime. Higher-order synergistic co-modulations carrying information about IG were found up to order 4. Although the largest percentage of significant links occurred at the pairwise level (545), significant triplets and quadruplets accounted for 14% and 2% of the total, respectively. In addition, we found that five triplets and the two quadruplets had a non-simplician structure, therefore suggesting a stronger emergent form of collective behaviour. The analysis of centrality of individual brain areas in higher-order behaviours showed that visual areas dominated pairwise functional interactions, whereas the vmPFC-OFC dominated both triplet and quadruplet interactions (Fig. 7B). In addition, feature-specific information transfer (FIT) analyses (Celotto *et al.*, 2023) showed that IG signals are broadcasted from multiple brain regions and converge to the ventromedial PFC and OFC (Fig. 8). These results further support the above-mentioned conclusion highlighting the central role of the vmPFC-OFC region in the processing of information gain during learning.

Generally, it is unclear how higher-order synergistic interactions arise from structural connectivity, and of how they impact the broadcasting of information relevant for cognition. From the theoretical point of view, further work is required to appreciate the relation between emergent higher-order behaviours, structural interactions and information dynamics. For example, an open question regards the relation between collective higher-order interdependencies as studied in the current work and structural higher-order interactions (Rosas *et al.*, 2022). Whereas the first pertains to the realm of what the system "does", the latter concerns how the system is structured. Completely disambiguating the causal structure exclusively from observational data is probably unrealistic. However, a proper inferential framework testing possible generative models of causal connections compatible with the data provides a viable solution (Friston, Harrison and Penny, 2003; Peel, Peixoto and De Domenico, 2022; Ramezanian-Panahi *et al.*, 2022; Zamora-López and Gilson, 2024). At the system neuroscience level, recent work have observed higher-order correlations beyond the pairwise in resting state EEG and fMRI activity (Gatica *et al.*, 2021; Herzog *et al.*, 2022; Santoro *et al.*, 2023; Hindriks *et al.*, 2024), which are largely invisible to functional connectivity and can display structured temporal patterns shifting from being redundancy-dominated to synergy-dominated (Varley, Pope,

Maria Grazia, *et al.*, 2023). We generalised this finding to cognitive settings and showed that higher-order interactions carry information in a synergistic manner and encode learning signals.

To conclude, the current study provides evidence that information gain is encoded in synergistic and higher-order functional interactions alongside with its broadcasting towards the prefrontal reward circuitry. Furthermore, it provides a novel perspective regarding how information relevant for cognition is encoded and broadcasted in distributed cortical networks and brain-wide dynamics.

# MATERIALS AND METHODS

### *Experimental conditions and behavioural learning task*

Eleven healthy subjects participated in the study (all right handed, 7 females; average age 26 years old). All participants gave written informed consent according to established institutional guidelines, and they received monetary compensation (50 euros) for their participation. The project has been approved by the local ethics committee.

Goal-directed learning was studied using an arbitrary visuomotor learning task design, where the relation between the visual stimulus, the action and its outcome is arbitrary and causal. The task required subjects to find by trial-and-error the correct associations between 3 coloured circles and 5 finger movements (Fig. 1a). On each trial, subjects were presented a coloured circle to which they had to respond within 1.5 s. Reaction times were computed as the time difference between stimulus presentation and motor response (finger movement). After a fixed delay of 1s following the disappearance of the digit number, an outcome image was presented for 1s and informed the subject whether the response was correct, incorrect, or too late (if the reaction time exceeded 1.5s). "Late" trials were excluded from the analysis, because they were either absent or very rare (i.e., maximum 2 late trials per session). The next trial started after a variable delay ranging from 2 to 3 s (randomly drawn from a uniform distribution) with the presentation of another visual stimulus. Visual stimuli were randomised in blocks of three trials. Each learning session was composed of 60 trials, 3 stimulus types (i.e., different colours, S1, S2, and S3) and 5 possible finger movements. Each subject performed 4 learning sessions, each containing different sets of coloured stimuli. To ensure reproducible performances across sessions and subjects, we developed a task design that manipulates learning and produced reproducible phases of acquisition and consolidation across sessions and individuals (Brovelli *et al.*, 2008). Anatomical T1-weighted MRI images were acquired for all participants using a 3-T whole-body imager equipped with a circular polarised head coil (Bruker). Magnetoencephalographic (MEG) recordings were performed using a 248 magnetometers system (4D Neuroimaging magnes 3600). Visual stimuli were projected using a video projection and motor responses were acquired using a LUMItouch® optical

response keypad with five keys. Presentation® software was used for stimulus delivery and experimental control during MEG acquisition.

### *MarsAtlas cortical parcellation and source model*

Single-subject cortical parcellation was performed using the MarsAtlas approach (Auzias, Coulon and Brovelli, 2016). After denoising using non-local means (Coupe *et al.*, 2008), T1-weighted MR-images were segmented using the FreeSurfer "recon-all" pipeline (v5.3.0). Grey and white matter segmentations were imported into the BrainVisa software (v4.5) and processed using the Morphologist pipeline procedure (http://brainvisa.info). White matter and pial surfaces were reconstructed and triangulated, and all sulci were detected and labelled automatically (Mangin et al., 2004). A parameterization of each hemisphere white matter mesh was performed using the Cortical Surface toolbox (http://www.meca-brain.org/softwares/). It resulted in a 2D orthogonal system defined on the white matter mesh, constrained by a set of primary and secondary sulci (Auzias et al., 2016). The MarsAtlas parcelization contained a total of 82 cortical parcels (41 each hemisphere). Each individual white matter mesh was then resampled onto a template defined by the average geometry across a population of 138 healthy subjects (https://meca-brain.org/software/hiphop138/) in two steps: 1) the parameterization induced by the 2D orthogonal system is used to map the white matter mesh onto a canonical spherical domain; 2) the location of each vertex of the template mesh is interpolated onto the individual mesh geometry using the classical barycentric mapping. This procedure defines one-to-one spatial correspondence across the white matter meshes from all individuals at the vertex level. By adapting the spatial resolution of the template mesh, we can fix the number of points that define the correspondence across subjects. The sources are then defined at the location of these points and are thus distributed at consistent anatomical locations across subjects.

Finally, we created source space and forward models readable by MNE-python (Gramfort *et al.*, 2013) using cortical meshes generated using BrainVISA (MarsAtlas) and subcortical structures generated by Freesurfer. These three elements are needed for the power estimation at the source level, which will be discussed in the next paragraph. These steps were performed using the BV2MNE toolbox (https://github.com/brainets/bv2mne), a python library developed in our team based on MNE-python . The anatomical meshes were coregistred with the MEG data using the 'mne coreg' interface.

### *Single-trial high-gamma activity (HGA) in MarsAtlas*

MEG signals were high pass filtered to 1Hz, low-pass filtered to 250 Hz, notch filtered in multiples of 50Hz and segmented into epochs aligned on stimulus and outcome presentation. Independent Component Analysis (ICA) was performed to detect and reject cardiac, eye-blink and oculomotor artefacts. Artefact rejection was performed semi automatically, at first by a visual inspection of the epochs' time series, then by means of the *autoreject* python library (Jas *et al.*, 2017) to detect and reject bad epochs from further analysis.

Spectral density estimation was performed using a multi-taper method based on discrete prolate spheroidal (slepian) sequences (Percival and Walden, 1993). To extract high-gamma activity from 60 to 120Hz, MEG time series were multiplied by *k* orthogonal tapers (*k* = 11) (0.2s in duration and 60Hz of frequency resolution, each stepped every 0.005s), centred at 90Hz and Fourier-transformed. Complex-valued estimates of spectral measures $X^{n}_{sensor}(t, k)$, including cross-spectral density matrices, were computed at the sensor level for each trial *n*, time *t* and taper *k*. Source analysis requires a physical forward model or leadfield, which describes the electromagnetic relation between sources and MEG sensors. The leadfield combines the geometrical relation of sources (dipoles) and sensors with a model of the conductive medium (i.e., the head model). For each participant, we generated a boundary element model using a single-shell model constructed from the segmentation of the cortical tissue obtained from individual MRI scans (Nolte, 2003). Lead-fields were not normalised. The head model, source locations and the information about MEG sensor position for both models were combined to derive single-participant leadfields. The orientation of cortical sources was set perpendicular to the cortical surface, whereas the orientation for subcortical sources was left unconstrained.

We used adaptive linear spatial filtering (Van Veen et al., 1997) to estimate the power at the source level. In particular, we employed the Dynamical Imaging of Coherent Sources (DICS) method, a beamforming algorithm for the tomographic mapping in the frequency domain (Gross et al., 2001), which is a well suited for the study of neural oscillatory responses based on single-trial source estimates of band-limited MEG signals (Hansen et al., 2015). At each source location, DICS employs a spatial filter that passes activity from this location with unit gain while maximally suppressing any other activity. The spatial

filters were computed on all trials for each time point and session, and then applied to single-trial MEG data.

Single-trial power estimates aligned with outcome and stimulus onset were log-transformed to make the data approximate Gaussian and low-pass filtered at 50Hz to reduce noise. Single-trial mean power and standard deviation in a time window from -0.5 and -0.1 s prior to stimulus onset was computed for each source and trial, and used to z-transform single-trial movement-locked power time courses. Similarly, single-trial outcome-locked power time courses were log-transformed and z-scored with respect to baseline period, to produce HGAs for the prestimulus period from -1.6 to -0.1 s with respect to stimulation for subsequent functional connectivity analysis. Finally, single-trial HGA for each brain region of *MarsAtlas* was computed as the mean z-transformed power values averaged across all sources within the same region. Single-trial HGA estimates were performed using home-made scripts based on MNE-python (Gramfort, 2013).

### *Model-based analysis of behavioural data and estimate of learning signals*

We estimated the evolution of stimulus-action-outcome probabilities and updating learning signals during learning using a Q-learning model (Watkins and Dayan, 1992) from reinforcement learning theory (Sutton and Barto, 1998). Briefly, the Q-learning model updates action values through the Rescorla-Wagner learning rule (1972) expressed by the following equation:

$$Q_{s,a}(t+1) = Q_{s,a}(t) + \alpha \cdot (r(t) - Q_a(t)) \quad (1)$$

where $Q_{s,a}(t)$ corresponds to stimulus-action value with s = 1, 2 or 3 (three possible coloured cues) and a = 1 to 5 (five possible finger movements), $r$ models the type of outcome (1 or 0 for correct and incorrect feedback) at trial *t,* and $\alpha$ is the learning rate (which varies from 0 to 1). Action values were then transformed into probabilities according to the softmax equation :

$$P_{s,a}(t) = exp(\beta Q_{s,a}(t)) \,/\, \sum_a exp(\beta Q_{s,a}(t)) \quad (2)$$

The coefficient β is termed the inverse 'temperature': low β (less than 1) causes all actions to be (nearly) equiprobable, whereas high β (greater than 1) amplifies the differences in association values. We identified the set of parameters that best fitted the behavioural data using a maximum likelihood approach. The model was fitted separately for each block of trials and learning session. For each learning session, we varied the learning rate

λ from 0.1 to 1 (in steps of 0.01) and β was varied from 1 to 10 (in steps of 0.2), and chose the set of parameters that maximise the log-likelihood of the probability to make the action performed by the participant $P_{chosen}$ defined as $L = \sum_{t} ln\, P_{chosen}(t)$.

Therefore, the Q-learning model provides estimates of the evolution of reward prediction-errors during learning as:

$$RPE(t) = \alpha \cdot (r(t) - Q_{s,a}(t)) \tag{3}$$

Bayesian inference is a principled statistical model based on Bayesian updating of beliefs about the current state of the world. The notion of information gain plays an important role in Bayesian update. In particular, information gain has also been defined via the change in models parameters during belief updating. This measure can be formalised via the Kullback–Leibler (KL) divergence between the prior and posterior, similarly defined as the information gain (Itti and Baldi, 2009; Modirshanechi *et al.*, 2023):

$$IG(t) = D_{KL}\left[P_{s,a}(t-1) \mid \mid P_{s,a}(t)\right] \tag{4}$$

Such information gain measures the belief-update upon a new outcome. Bayesian surprise measures the pseudo-distance between the belief of obtaining a correct outcome for a given stimulus-action pair prior to the outcome $P_{s,a}(t-1)$and after the outcome $P_{s,a}(t)$. This means that an event that causes a bigger change in the model of the world is more surprising. In this context, IG quantifies the amount of information gain or model update at each trial, that is after the observation of action outcome.

### *Cortical areas encoding learning signals*

We used information-theoretic metrics to quantify the statistical dependency between single-trial HGA and the outcome-related learning signals. Information-based measures quantify how much the neural activity of a single brain region explains a variable of the task. To this end, we computed the mutual information (MI); as a reminder, mutual information is defined as:

$$I(X;Y) = H(X) - H(X|Y) \tag{5}$$

In this equation the variables $X$ and $Y$ represent the HGA power and the behavioural variables, respectively. $H(X)$ is the entropy of $X$, and $H(X|Y)$ is the conditional entropy of $X$ given $Y$. In particular, we used Gaussian-Copula Mutual Information (GCMI) (Ince et al.,

2017), which is a semi-parametric binning-free technique to calculate MI. The GCMI is a robust rank-based approach that allows detecting any type of monotonic relation between variables. More precisely, the GCMI is a lower-bound approximate estimator of mutual information for continuous signals based on a Gaussian copula, and it is of practical importance for the analysis of short and potentially multivariate neural signals. Note, however, that the GCMI does not detect non-monotonic (e.g., parabolic) relations. In the current work, the GCMI was computed across trials between time-resolved HGA and outcome-related learning signals. The considered outcome-related learning signals were reward prediction errors (RPE) and Information Gain (IG).

### *Cortico-cortical functional interactions encoding learning signals*

The goal of network-level analyses was to identify dynamic functional networks encoding outcome-related computations, such as the reward prediction errors (RPE) and information gain (IG). We used information-theoretic metrics to quantify the statistical dependency between single-trial event-related co-modulations in HGA across trials and learning computations. In general terms, the aim of our analyses was to assess the interdependence between pairs of brain signals ($X_1$ and $X_2$) and a third learning variable (Y), either RPE or IG. More precisely, we wished to quantify whether the interaction in activity co-modulation between brain regions supported functional segregation and integration phenomena supporting learning. We reasoned that functional segregation could be quantified as the amount of redundant information shared in the co-modulation of activity across brain regions. Similarly, we propose that functional integration processes could be quantified as the amount of synergistic information carried by the two brain regions about the target learning variable. Such research questions can be formalised within the Partial Information Decomposition (PID) framework (Williams and Beer, 2010; Lizier *et al.*, 2018). Indeed, the PID framework allows the decomposition of multivariate mutual information between a system of predictors and a target variable, and to quantify the information that several sources (or predictors) variables provide *uniquely*, *redundantly* or *synergistically* about a target variable. In the current study, we considered the case where the source variables are the across-trials HGA of two brain regions ($X_1$ and $X_2$) and the target variable is the trial-by-trial evolution of the learning variable (Y). If pairs of brain regions $X_1$ and $X_2$ carry the same information about learning variable Y, we refer to it as *redundant* information. If only one area $X_1$ or $X_2$ carries information about Y, it is referred to

as *unique* information carried by $X_1$ or $X_2$ about Y, and vice-versa for Y. Finally, if neither of the variables alone provide information about Y, but they need to be observed together, we refer to it as *synergistic* information. In other words, the knowledge of any of the predictors separately does not provide any information about the target variable C. Analytically, the PID proposes to decompose the total mutual information between a pair of source variables X and Y and a target variable Z $I(X_1, X_2; Y)$ into four non-negative components:

$$I(X_1, X_2; Y) = U(X_1; Y) + U(X_2; Y) + R(X_1, X_2; Y) + S(X_1, X_2; Y) \quad (6)$$

In addition, the PID formulation links to classical Shannon measures of mutual information as

$$I(X_1; Y) = U(X_1; Y) + R(X_1, X_2; Y) \quad (7)$$

$$I(X_2; Y) = U(X_2; Y) + R(X_1, X_2; Y) \quad (8)$$

The problem with this approach is that its governing equations form an underdetermined system: only three quantities that can be computed from the data (i.e., the mutual information quantities $I(X_1, X_2; Y)$, $I(X_1; Y)$ and $I(X_2; Y)$) for the four terms of the PID (i.e., two unique information, redundant and synergistic). To actually calculate the decomposition, an additional assumption must therefore be made. Here, we exploit the so-called minimum mutual information (MMI) PID, which has been shown to provide correct estimations for a broad class of systems following a multivariate Gaussian distribution (Barrett, 2015; Luppi *et al.*, 2022). According to the MMI PID, redundant information carried by pairs of brain regions is given by the minimum of the information provided by each individual source to the target,

$$R_{MMI}(X_1, X_2; Y) = min[I(X_1; Y), I(X_2; Y)] \quad (9)$$

Then, synergistic information can be computed by substituting Eq. 8, 9 and 10 in Eq. 7 and rearranging the terms

$$S_{MMI}(X_1, X_2; Y) = I(X_1, X_2; Y) - I(X_1; Y) - I(X_2; Y) + R_{MMI}(X_1, X_2; Y) \quad (10)$$

Equations 9 and 10 represent the redundant and synergistic information carried by the co-modulations in HGA of pairs of brain regions $X_1$ and $X_2$ about the learning variable Z, respectively. We will refer to these quantities as redundant (RC) and synergistic connectivity (SC), respectively. By definition, redundant functional interactions carry the same information about learning variables than the one carried by individual brain areas. On the other hand, synergistic connectivity reveals brain functional interactions that cannot

be explained by individual contribution, but only by their combined and collective coordination.

### *Higher-order cortico-cortical interactions encoding learning signals*

We then studied whether outcome-related learning signals would be encoded in the interactions beyond the pairwise relations, the so-called higher-order functional interactions, or correlations. The definitions of redundancy and synergy can be generalised to higher-order by simply iterating over all regions of a multiple. By considering a multiple as a set of $n$ variables $X^n = \{X_1, \ldots, X_n\}$, the redundant information carried by each higher-order multiplet about the target variable is given by:

$$R_{MMI}(X^n; Y) = min_i \, I(X_i; Y) \tag{11}$$

Then, synergistic information can be computed by substituting Eq. 8, 9 and 10 in Eq. 7 and rearranging the terms

$$S_{MMI}(X^n; Y) = I(X^n; Y) - max_i \, I(X_i^-; Y) \tag{12}$$

where $X_i^-$ is the the set of variables of $X^n$ excluding the brain area $i$. One of the main issues in the study of the higher-order behaviours of complex systems is the computational cost required to investigate all the multiplets of any order. In order to perform HOIs analyses over all orders, we thus identified 9 sets of brain regions that displayed significant encoding of learning signals and we averaged the HGA for each set of regions-of-interest. This allowed us to perform an exhaustive analysis of HOIs at all orders. Equations 11 and 12 represent the redundant and synergistic information carried by multiplets of brain regions about the learning variable Y, respectively. We will refer to these quantities as higher-order redundant (HiRFC) and synergistic connectivity (HiSFC), respectively.

### *Cortico-cortical information transfer encoding learning signals*

In order to investigate whether learning signals are broadcasted across different cortical regions, we used a recently-developed information-theoretic measure termed Feature-specific Information Transfer (FIT) that quantifies how much information about specific features flows between brain areas (Celotto *et al.*, 2023). FIT merges the Wiener-Granger causality principle (Granger, 1980; Brovelli *et al.*, 2004; Bressler and Seth,

2011) with content specificity based on the PID framework (Williams and Beer, 2010; Lizier *et al.*, 2018). FIT isolates information about a specific task variable Y encoded in the current activity of a receiving neural population, which was not encoded in its past activity, and which was instead encoded by the past activity of the sender neural population. More precisely, the FIT measure is based on two four-variables PID and it is defined as the minimum of two "atoms". The first atom is defined in a four-variable PID, which considers the learning variable $Y$ as the target and three source variables, namely the brain activity in the past of the sending signal $X_1(t-\tau)$, the past of the receiver $X_2(t-\tau)$ and the present of the receiver $X_2(t)$. This atom formalises the principle of Granger-Wiener causality within the PID framework and it adds the encoding of task variables (e.g., learning) in the time-lagged predictability between signals. More formally, the first atom is defined as the redundant information about $Y$ carried by the past of $X_1(t-\tau)$ and the present of $X_2(t)$, which is unique with respect to past of the $X_2(t-\tau)$. As for previous analyses, we used the MMI PID to compute this atom, which was computed as

$$R_{MMI}(X_1(t-\tau),\, X_2(t); Y) = min[I(X_1(t-\tau);\, Y),\, I(X_2(t);\, Y)] - min([I(X_1(t-\tau);\, Y),\, I(X_2(t);\, Y),\, I(X_2(t-\tau);\, Y)] \quad (13)$$

The second term on the right-hand side assures the uniqueness with respect to the past of the receiver $(X_2(t-\tau)$. The second atom for the calculation of the FIT is defined for a four-variable PID, considering the present of the receiver $X_2(t)$, as the target variable and the learning variable $Y$, $X_1(t-\tau)$ and $X_2(t-\tau)$ as source variables. This second term assures that the FIT measure does exceed: i) the overall propagation of information between signals, referred to as the Directed Information (Massey 1990) or Transfer Entropy (Schreiber 2000), and quantified by means of the conditional mutual information $I(X_1(t-\tau);\, X_2(t)|X_2(t-\tau))$; ii) the mutual information between the learning variable and the past of the sender, $I(Y; X_1(t-\tau))$ and iii) the mutual information between the learning variable and the present of the receiver, $I(Y; X_2(t))$. This term is defined as the redundant information about the present of the receiver $X_2(t)$ carried by $X_1(t-\tau)$ and $Y$, which is unique with respect to the past of the receiver $X_2(t-\tau)$:

$$R_{MMI}(X_1(t-\tau),\, Y; X_2(t)) = min[I(X_1(t-\tau);\, X_2(t)),\, I(Y,\, X_2(t))]$$

$$- min[I(X_1(t - \tau);\ X_2(t)),\ I(Y,\ X_2(t)),\ I(X_2(t - \tau);\ X_2(t)] \quad (14)$$

Therefore, the FIT measure is defined as the minimum between these two atoms (Eq. 13 and 14) to select the smallest nonnegative piece of information:

$$FIT\ =\ min[R_{MMI}(X_1(t - \tau),\ X_2(t); Y)\ ,\ R_{MMI}(X_1(t - \tau),\ Y; X_2(t))\ ] \quad (15)$$

We computed the FIT measure between all pairs of 9 sets of brain regions at different delays tau ranging from -0.005 to -0.2s, and averaged over delays. Statistical analysis was similarly applied as for all other measures.

### *Statistical analysis*

For the statistical inferences, we used a group-level approach based on non-parametric permutations and encompassing non-negative measures of information (Combrisson, Allegra, *et al.*, 2022). We used a random effect (RFX) to take into account the inter-subject variability. In this approach, the information theoretical metrics (e.g., mutual information, MI) between the neurophysiological signal and the behavioural regressor are computed across trials for each participant separately, at each time point and brain region. To sample the distribution of MI attainable by chance, we computed the MI between the brain data and a randomly shuffled version of the behavioural variable. This procedure was then repeated 1000 times. We took the mean of the MI values computed on the permutations, and used this mean (MI) to perform a one sample t-test across all the participants' MI values obtained both from original and permuted data. We then used a cluster-based approach to assess whether the size of the estimated t-values significantly differs from its distribution. The cluster forming threshold was defined as the 95th percentile of the distribution of t-values. We used this cluster forming threshold to identify the cluster mass of t-values on both original and permuted data. Finally, to correct for multiple comparisons across both time and space, we build a distribution made of the 1000 largest clusters estimated on the permuted data. The final corrected p-values were inferred as the proportion of permutations exceeding the t-values. For what concerns functional connectivity analysis, the information theoretical metrics of interest (e.g., redundant connectivity) across pairs of brain regions were studied similarly as the local MI time courses.

### *Code*

Higher-order redundancy and synergistic measures were computed using the HOI

toolbox (https://github.com/brainets/hoi). All pair-wise information theoretical measures, including the FIT, and statistical analyses were performed using the Frites Python package (https://github.com/brainets/frites) (Combrisson, Basanisi, *et al.*, 2022).

# References


Allen, W.E. *et al.* (2019) 'Thirst regulates motivated behavior through modulation of brainwide neural population dynamics', *Science*, 364(6437), p. 253.

Auzias, G., Coulon, O. and Brovelli, A. (2016) 'MarsAtlas: A cortical parcellation atlas for functional mapping', *Human brain mapping*, 37(4), pp. 1573–1592.

Averbeck, B.B. and Costa, V.D. (2017) 'Motivational neural circuits underlying reinforcement learning', *Nature neuroscience*, 20(4), pp. 505–512.

Averbeck B, M.E.A. (2020) 'Hypothalamic Interactions with Large-Scale Neural Circuits Underlying Reinforcement Learning and Motivated Behavior', *Trends Neurosci* 43(9):681-694.

Averbeck, B. and O'Doherty, J.P. (2021) 'Reinforcement-learning in fronto-striatal circuits', *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 47(1), pp. 147–162.

Badre, D. *et al.* (2012) 'Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration', *Neuron*, 73(3), pp. 595–607.

Badre, D. and D'Esposito, M. (2009) 'Is the rostro-caudal axis of the frontal lobe hierarchical?', *Nature reviews. Neuroscience*, 10(9), pp. 659–669.

Baldi, P. and Itti, L. (2010) 'Of bits and wows: A Bayesian theory of surprise with applications to attention', *Neural networks: the official journal of the International Neural Network Society*, 23(5), pp. 649–666.

Balleine, B.W. (2019) 'The Meaning of Behavior: Discriminating Reflex and Volition in the Brain', *Neuron*, 104(1), pp. 47–62.

Balleine, B.W. and O'Doherty, J.P. (2010) 'Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action', *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 35(1), pp. 48–69.

Ballesta, S. *et al.* (2020) 'Values encoded in orbitofrontal cortex are causally related to economic choices', *Nature*, 588(7838), pp. 450–453.

Bandura, A. (1997) 'Self-efficacy: The exercise of control', 604. Available at: https://psycnet.apa.org/fulltext/1997-08589-000.pdf.

Barrett, A.B. (2015) 'Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems', *Physical review. E, Statistical, nonlinear, and soft matter physics*, 91(5), p. 052802.

Bartolo, R. and Averbeck, B.B. (2020) 'Prefrontal Cortex Predicts State Switches during Reversal Learning', *Neuron*, 106(6), pp. 1044–1054.e4.

Bassett, D.S. and Mattar, M.G. (2017) 'A Network Neuroscience of Human Learning: Potential to Inform Quantitative Theories of Brain and Behavior', *Trends in cognitive sciences*, 21(4), pp. 250–264.

Battaglia, D. and Brovelli, A. (2020) 'Functional connectivity and neuronal dynamics: insights from computational methods'. The Cognitive Neurosciences (Sixth Edition), Edited by David Poeppel, George R. Mangun, Michael S. Gazzaniga.

Battiston, F. *et al.* (2020) 'Networks beyond pairwise interactions: Structure and dynamics', *Physics reports*, 874, pp. 1–92.

Battiston, F. *et al.* (2021) 'The physics of higher-order interactions in complex systems', *Nature physics*, 17(10), pp. 1093–1098.

Benson, A.R., Gleich, D.F. and Leskovec, J. (2016) 'Higher-order organization of complex networks', *Science*, 353(6295), pp. 163–166.

Braun, U. *et al.* (2015) 'Dynamic reconfiguration of frontal brain networks during executive cognition in humans', *Proceedings of the National Academy of Sciences of the United States of America*, 112(37), pp. 11678–11683.

Bressler, S.L. and Menon, V. (2010) 'Large-scale brain networks in cognition: emerging methods and principles', *Trends in cognitive sciences*, 14(6), pp. 277–290.

Bressler, S.L. and Seth, A.K. (2011) 'Wiener–Granger Causality: A well established methodology', *NeuroImage*, 58(2), pp. 323–329.

Bromberg-Martin, E.S. *et al.* (2024) 'A neural mechanism for conserved value computations integrating information and rewards', *Nature neuroscience*, 27(1), pp. 159–175.

Bromberg-Martin, E.S. and Hikosaka, O. (2009) 'Midbrain dopamine neurons signal preference for advance information about upcoming rewards', *Neuron*, 63(1), pp. 119–126.

Bromberg-Martin, E.S. and Hikosaka, O. (2011) 'Lateral habenula neurons signal errors in the prediction of reward information', *Nature neuroscience*, 14(9), pp. 1209–1216.

Bromberg-Martin, E.S., Matsumoto, M. and Hikosaka, O. (2010) 'Dopamine in motivational control: rewarding, aversive, and alerting', *Neuron*, 68(5), pp. 815–834.

Brovelli, A. *et al.* (2004) 'Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality', *Proceedings of the National Academy of Sciences of the United States of America*, 101(26), pp. 9849–9854.

Brovelli, A. *et al.* (2008) 'Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory', *Cerebral cortex* , 18(7), pp. 1485–1495.

Brydevall, M. *et al.* (2018) 'The neural encoding of information prediction errors during non-instrumental information seeking', *Scientific reports*, 8(1), p. 6134.

Buehlmann, A. and Deco, G. (2010) 'Optimal information transfer in the cortex through

synchronization', *PLoS computational biology*, 6(9).

Buzsáki, G. and Draguhn, A. (2004) 'Neuronal oscillations in cortical networks', *Science*, 304(5679), pp. 1926–1929.

Cavanagh, J.F. *et al.* (2012) 'Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation', *Cerebral cortex* , 22(11), pp. 2575–2586.

Cavanagh, J.F. and Frank, M.J. (2014) 'Frontal theta as a mechanism for cognitive control', *Trends in cognitive sciences*, 18(8), pp. 414–421.

Celotto, M. *et al.* (2023) 'An information-theoretic quantification of the content of communication between brain regions', Advances in Neural Information Processing Systems (NeurIPS) 36, 36, 64213–64265

Charpentier, C.J., Bromberg-Martin, E.S. and Sharot, T. (2018) 'Valuation of knowledge and ignorance in mesolimbic reward circuitry', *Proceedings of the National Academy of Sciences of the United States of America*, 115(31), pp. E7255–E7264.

Chelaru, M.I. *et al.* (2021) 'High-order interactions explain the collective behavior of cortical populations in executive but not sensory areas', *Neuron*, 109(24), pp. 3954–3961.e5.

Chen, S. *et al.* (2024) 'Brain-wide neural activity underlying memory-guided movement', *Cell*, 187(3), pp. 676–691.e16.

Cockburn, J. *et al.* (2022) 'Novelty and uncertainty regulate the balance between exploration and exploitation through distinct mechanisms in the human brain', *Neuron*, 110(16), pp. 2691–2702.e8.

Cohen, J.D., McClure, S.M. and Yu, A.J. (2007) 'Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1481), pp. 933–942.

Cohen, J.R. and D'Esposito, M. (2016) 'The Segregation and Integration of Distinct Brain Networks and Their Relationship to Cognition', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 36(48), pp. 12083–12094.

Combrisson, E., Basanisi, R., *et al.* (2022) 'Frites: A Python package for functional connectivity analysis and group-level statistics of neurophysiological data', *Journal of open source software*, 7(79), p. 3842.

Combrisson, E., Allegra, M., *et al.* (2022) 'Group-level inference of information-based measures for the analyses of cognitive brain networks from neurophysiological data', *NeuroImage*, 258, p. 119347.

Combrisson, E. *et al.* (2023) 'Neural interactions in the human frontal cortex dissociate reward and punishment learning', *eLife*, 12.

Corbetta, M. and Shulman, G.L. (2011) 'Spatial neglect and attention networks', *Annual review of neuroscience*, 34, pp. 569–599.

Coupe, P. *et al.* (2008) 'An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images', *IEEE transactions on medical imaging*, 27(4), pp. 425–441.

Courville, A.C., Daw, N.D. and Touretzky, D.S. (2006) 'Bayesian theories of conditioning in a changing world', *Trends in cognitive sciences*, 10(7), pp. 294–300.

Crutchfield, J.P. (1994) 'The calculi of emergence: computation, dynamics and induction', *Physica D. Nonlinear phenomena*, 75(1), pp. 11–54.

D'Ardenne, K. *et al.* (2008) 'BOLD responses reflecting dopaminergic signals in the human ventral tegmental area', *Science*, 319(5867), pp. 1264–1267.

Daw, N.D., Niv, Y. and Dayan, P. (2005) 'Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control', *Nature neuroscience*, 8(12), pp. 1704–1711.

Deco, G. *et al.* (2015) 'Rethinking segregation and integration: contributions of whole-brain modelling', *Nature reviews. Neuroscience*, 16(7), pp. 430–439.

Dickinson, A. (1994) 'CHAPTER 3 - Instrumental Conditioning', in N.J. Mackintosh (ed.) *Animal Learning and Cognition*. San Diego: Academic Press, pp. 45–79.

Dolan, R.J. and Dayan, P. (2013) 'Goals and habits in the brain', *Neuron*, 80(2), pp. 312–325.

Engel, A.K., Fries, P. and Singer, W. (2001) 'Dynamic predictions: Oscillations and synchrony in top–down processing', *Nature reviews. Neuroscience*, 2(10), pp. 704–716.

Faraji, M., Preuschoff, K. and Gerstner, W. (2018) 'Balancing New against Old Information: The Role of Puzzlement Surprise in Learning', *Neural computation*, 30(1), pp. 34–83.

Finc, K. *et al.* (2020) 'Dynamic reconfiguration of functional brain networks during working memory training', *Nature communications*, 11(1), p. 2435.

Fouragnan, E., Retzler, C. and Philiastides, M.G. (2018) 'Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis', *Human brain mapping*, 39(7), pp. 2887–2906.

Frank, M.J. *et al.* (2009) 'Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation', *Nature neuroscience*, 12(8), pp. 1062–1068.

Frank, M.J., Seeberger, L.C. and O'reilly, R.C. (2004) 'By carrot or by stick: cognitive reinforcement learning in parkinsonism', *Science*, 306(5703), pp. 1940–1943.

Fries, P. (2015) 'Rhythms for Cognition: Communication through Coherence', *Neuron*, 88(1), pp. 220–235.

Friston, K. *et al.* (2015) 'Active inference and epistemic value', *Cognitive neuroscience*, 6(4), pp. 187–214.

Friston, K. *et al.* (2017) 'Active Inference: A Process Theory', *Neural computation*, 29(1), pp. 1–49.

Friston, K.J., Harrison, L. and Penny, W. (2003) ‘Dynamic causal modelling’, *NeuroImage*, 19(4), pp. 1273–1302.

Ganmor, E., Segev, R. and Schneidman, E. (2011) ‘Sparse low-order interaction network underlies a highly correlated and learnable neural population code’, *Proceedings of the National Academy of Sciences of the United States of America*, 108(23), pp. 9679–9684.

Gardner, M.P.H. *et al.* (2020) ‘Processing in Lateral Orbitofrontal Cortex Is Required to Estimate Subjective Preference during Initial, but Not Established, Economic Choice’, *Neuron*, 108(3), pp. 526–537.e4.

Gatica, M. *et al.* (2021) ‘High-Order Interdependencies in the Aging Brain’, *Brain connectivity*, 11(9), pp. 734–744.

Gershman, S.J. (2018) ‘Deconstructing the human algorithms for exploration’, *Cognition*, 173, pp. 34–42.

Gläscher, J. *et al.* (2010) ‘States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning’, *Neuron*, 66(4), pp. 585–595.

Gore, F. *et al.* (2023) ‘Orbitofrontal cortex control of striatum leads economic decision-making’, *Nature neuroscience*, 26(9), pp. 1566–1574.

Gottlieb, J. *et al.* (2013) ‘Information-seeking, curiosity, and attention: computational and neural mechanisms’, *Trends in cognitive sciences*, 17(11), pp. 585–593.

Gottlieb, J. and Oudeyer, P.-Y. (2018) ‘Towards a neuroscience of active sampling and curiosity’, *Nature reviews. Neuroscience*, 19(12), pp. 758–770.

Gramfort, A. *et al.* (2013) ‘MEG and EEG data analysis with MNE-Python’, *Frontiers in neuroscience*, 7, p. 267.

Granger, C.W.J. (1980) ‘Testing for causality: A personal viewpoint’, *Journal of economic dynamics & control*, 2, pp. 329–352.

Grilli, J. *et al.* (2017) ‘Higher-order interactions stabilize dynamics in competitive network models’, *Nature*, 548(7666), pp. 210–213.

Gueguen, M.C.M. *et al.* (2021) ‘Anatomical dissociation of intracerebral signals for reward and punishment prediction errors in humans’, *Nature communications*, 12(1), p. 3344.

Haber, S.N. and Knutson, B. (2009) ‘The Reward Circuit: Linking Primate Anatomy and Human Imaging’, *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 35(1), pp. 4–26.

Herzog, R. *et al.* (2022) ‘Genuine high-order interactions in brain networks and neurodegeneration’, *Neurobiology of disease*, 175, p. 105918.

Hindriks, R. *et al.* (2024) ‘Higher-order functional connectivity analysis of resting-state functional magnetic resonance imaging data using multivariate cumulants’, *Human brain*

*mapping*, 45(5), p. e26663.

Hunt, L.T. and Hayden, B.Y. (2017) 'A distributed, hierarchical and recurrent framework for reward-based choice', *Nature reviews. Neuroscience*, 18(3), pp. 172–182.

International Brain Laboratory *et al.* (2023) 'A Brain-Wide Map of Neural Activity during Complex Behaviour', *bioRxiv*. Available at: https://doi.org/10.1101/2023.07.04.547681.

Itti, L. and Baldi, P. (2009) 'Bayesian surprise attracts human attention', *Vision research*, 49(10), pp. 1295–1306.

Jas, M. *et al.* (2017) 'Autoreject: Automated artifact rejection for MEG and EEG data', *NeuroImage*, 159, pp. 417–429.

Jocham, G. *et al.* (2016) 'Reward-Guided Learning with and without Causal Attribution', *Neuron*, 90(1), pp. 177–190.

Khilkevich, A. *et al.* (2024) 'Brain-wide dynamics linking sensation to action during decision-making', *Nature*, pp. 1–11.

Kirst, C., Timme, M. and Battaglia, D. (2016) 'Dynamic information routing in complex networks', *Nature communications*, 7, p. 11061.

Kobayashi, K. *et al.* (2021) 'Dynamic Representation of the Subjective Value of Information', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 41(39), pp. 8220–8232.

Kobayashi, K. and Hsu, M. (2019) 'Common neural code for reward and information value', *Proceedings of the National Academy of Sciences of the United States of America*, 116(26), pp. 13061–13066.

Kobayashi, K. and Kable, J.W. (2024) 'Neural mechanisms of information seeking', *Neuron*, 112(11), pp. 1741–1756.

Lee, S.W., Shimojo, S. and O'Doherty, J.P. (2014) 'Neural computations underlying arbitration between model-based and model-free learning', *Neuron*, 81(3), pp. 687–699.

Levine, J.M. *et al.* (2017) 'Beyond pairwise mechanisms of species coexistence in complex communities', *Nature*, 546(7656), pp. 56–64.

Levy, D.J. and Glimcher, P.W. (2012) 'The root of all value: a neural common currency for choice', *Current opinion in neurobiology*, 22(6), pp. 1027–1038.

Liakoni, V. *et al.* (2021) 'Learning in Volatile Environments With the Bayes Factor Surprise', *Neural computation*, 33(2), pp. 269–340.

Liakoni, V. *et al.* (2022) 'Brain signals of a Surprise-Actor-Critic model: Evidence for multiple learning modules in human decision making', *NeuroImage*, 246, p. 118780.

Liljeholm, M. *et al.* (2011) 'Neural correlates of instrumental contingency learning: differential effects of action-reward conjunction and disjunction', *The Journal of*

*neuroscience: the official journal of the Society for Neuroscience*, 31(7), pp. 2474–2480.

Liljeholm, M. *et al.* (2013) 'Neural correlates of the divergence of instrumental probability distributions', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 33(30), pp. 12519–12527.

Lizier, J.T. *et al.* (2018) 'Information Decomposition of Target Effects from Multi-Source Interactions: Perspectives on Previous, Current and Future Work', *Entropy* , 20(4).

Luppi, A.I. *et al.* (2022) 'A synergistic core for human brain evolution and cognition', *Nature neuroscience*, 25(6), pp. 771–782.

Luppi, A.I. *et al.* (2024) 'Information decomposition and the informational architecture of the brain', *Trends in cognitive sciences*, 28(4), pp. 352–368.

Mackintosh, N.J. (1975) 'A theory of attention: Variations in the associability of stimuli with reinforcement', *Psychological review*, 82(4), pp. 276–298.

Martignon, L. *et al.* (2000) 'Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies', *Neural computation*, 12(11), pp. 2621–2653.

Mediano, P.A.M. *et al.* (2022) 'Integrated information as a common signature of dynamical and information-processing complexity', *Chaos* , 32(1), p. 013115.

Mehlhorn, K. *et al.* (2015) 'Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures', *Decisions*, 2(3), pp. 191–215.

Modirshanechi, A. *et al.* (2023) 'Surprise and novelty in the brain', *Current opinion in neurobiology*, 82, p. 102758.

Modirshanechi, A., Brea, J. and Gerstner, W. (2022) 'A taxonomy of surprise definitions', *Journal of mathematical psychology*, 110, p. 102712.

Monosov, I.E. (2024) 'Curiosity: primate neural circuits for novelty and information seeking', *Nature reviews. Neuroscience*, 25(3), pp. 195–208.

Montani, F. *et al.* (2009) 'The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex', *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367(1901), pp. 3297–3310.

Morris, R.W. *et al.* (2022) 'The neural bases of action-outcome learning in humans', *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 42 (17) 3636-3647

Nigam, S., Pojoga, S. and Dragoi, V. (2019) 'Synergistic Coding of Visual Information in Columnar Networks', *Neuron*, 104(2), pp. 402–411.e4.

Norton, K.G. and Liljeholm, M. (2020) 'The Rostrolateral Prefrontal Cortex Mediates a Preference for High-Agency Environments', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 40(22), pp. 4401–4409.

Nour, M.M. *et al.* (2018) 'Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia', *Proceedings of the National Academy of Sciences of the United States of America*, 115(43), pp. E10167–E10176.

O'Doherty, J. *et al.* (2004) 'Dissociable roles of ventral and dorsal striatum in instrumental conditioning', *Science*, 304(5669), pp. 452–454.

Ohiorhenuan, I.E. *et al.* (2010) 'Sparse coding and high-order correlations in fine-scale cortical networks', *Nature*, 466(7306), pp. 617–621.

O'Reilly, J.X. *et al.* (2013) 'Dissociable effects of surprise and model update in parietal and anterior cingulate cortex', *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), pp. E3660–9.

Padoa-Schioppa, C. and Assad, J.A. (2006) 'Neurons in the orbitofrontal cortex encode economic value', *Nature*, 441(7090), pp. 223–226.

Palmigiano, A. *et al.* (2017) 'Flexible information routing by transient synchrony', *Nature neuroscience*, 20(7), pp. 1014–1022.

Palminteri, S. *et al.* (2015) 'Contextual modulation of value signals in reward and punishment learning', *Nature communications*, 6, p. 8096.

Panzeri, S. *et al.* (2022) 'The structures and functions of correlations in neural population codes', *Nature reviews. Neuroscience*, 23(9), pp. 551–567.

Pearce, J.M. and Hall, G. (1980) 'A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli', *Psychological review*, 87(6), pp. 532–552.

Peel, L., Peixoto, T.P. and De Domenico, M. (2022) 'Statistical inference links data and theory in network science', *Nature communications*, 13(1), p. 6794.

Percival, D.B. and Walden, A.T. (1993) *Spectral Analysis for Physical Applications*. Cambridge University Press.

Pessiglione, M. *et al.* (2006) 'Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans', *Nature*, 442(7106), pp. 1042–1045.

Proca, A.M. *et al.* (2024) 'Synergistic information supports modality integration and flexible learning in neural networks solving multiple tasks', *PLoS computational biology*, 20(6), p. e1012178.

Raja Beharelle, A. *et al.* (2015) 'Transcranial Stimulation over Frontopolar Cortex Elucidates the Choice Attributes and Neural Mechanisms Used to Resolve Exploration-Exploitation Trade-Offs', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 35(43), pp. 14544–14556.

Ramezanian-Panahi, M. *et al.* (2022) 'Generative Models of Brain Dynamics', *Frontiers in artificial intelligence*, 5, p. 807406.

Rescorla, R.A. (1991) 'Associative Relations in Instrumental Learning: The Eighteenth Bartlett Memorial Lecture', *The Quarterly Journal of Experimental Psychology Section B*, 43(1b), pp. 1–23.

Rosas, F.E. *et al.* (2020) 'An operational information decomposition via synergistic disclosure', *Journal of Physics A: Mathematical and Theoretical*, 53(48), p. 485001.

Rosas, F.E. *et al.* (2022) 'Disentangling high-order mechanisms and high-order behaviours in complex systems', *Nature physics*, 18(5), pp. 476–477.

Sanchez-Gorostiaga, A. *et al.* (2019) 'High-order interactions distort the functional landscape of microbial consortia', *PLoS biology*, 17(12), p. e3000550.

Santoro, A. *et al.* (2023) 'Higher-order organization of multivariate time series', *Nature physics*, 19(2), pp. 221–229.

Schneidman, E. *et al.* (2006) 'Weak pairwise correlations imply strongly correlated network states in a neural population', *Nature*, 440(7087), pp. 1007–1012.

Schultz, W. (2006) 'Behavioral theories and the neurophysiology of reward', *Annual review of psychology*, 57, pp. 87–115.

Schwartenbeck, P. *et al.* (2013) 'Exploration, novelty, surprise, and free energy minimization', *Frontiers in psychology*, 4, p. 710.

Schwartenbeck, P. *et al.* (2019) 'Computational mechanisms of curiosity and goal-directed exploration', *eLife*, 8.

Shahidi, N. *et al.* (2019) 'High-order coordination of cortical spiking activity modulates perceptual accuracy', *Nature neuroscience*, 22(7), pp. 1148–1158.

Shimazaki, H. *et al.* (2012) 'State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data', *PLoS computational biology*, 8(3), p. e1002385.

Shine, J.M. *et al.* (2016) 'The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance', *Neuron*, 92(2), pp. 544–554.

Shulman, G.L. *et al.* (2010) 'Right hemisphere dominance during spatial selective attention and target detection occurs outside the dorsal frontoparietal network', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 30(10), pp. 3640–3651.

Sizemore, A.E. *et al.* (2019) 'The importance of the whole: Topological data analysis for the network neuroscientist', *Network neuroscience (Cambridge, Mass.)*, 3(3), pp. 656–673.

Sporns, O. (2013) 'Network attributes for segregation and integration in the human brain', *Current opinion in neurobiology*, 23(2), pp. 162–171.

Steinmetz, N.A. *et al.* (2019) 'Distributed coding of choice, action and engagement across the mouse brain', *Nature*, 576(7786), pp. 266–273.

Stringer, C. *et al.* (2019) 'Spontaneous behaviors drive multidimensional, brainwide activity', *Science*, 364(6437), p. 255.

Sutton, R.S. and Barto, A.G. (2018) *Reinforcement Learning, second edition: An Introduction*. MIT Press.

Tanaka, S.C., Balleine, B.W. and O'Doherty, J.P. (2008) 'Calculating consequences: brain systems that encode the causal effects of actions', *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 28(26), pp. 6750–6755.

Tax, T.M.S., Mediano, P.A.M. and Shanahan, M. (2017) 'The Partial Information Decomposition of Generative Neural Network Models', *Entropy* , 19(9), p. 474.

Tishby, N. and Polani, D. (2011) 'Information Theory of Decisions and Actions', in V. Cutsuridis, A. Hussain, and J.G. Taylor (eds) *Perception-Action Cycle: Models, Architectures, and Hardware*. New York, NY: Springer New York, pp. 601–636.

Tolman, E.C. (1948) 'Cognitive maps in rats and men', *Psychological review*, 55(4), pp. 189–208.

Varela, F. *et al.* (2001) 'The brainweb: phase synchronization and large-scale integration', *Nature reviews. Neuroscience*, 2(4), pp. 229–239.

Varley, T.F., Sporns, O., *et al.* (2023) 'Information-processing dynamics in neural networks of macaque cerebral cortex reflect cognitive state and behavior', *Proceedings of the National Academy of Sciences of the United States of America*, 120(2), p. e2207677120.

Varley, T.F., Pope, M., Faskowitz, J., *et al.* (2023) 'Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex', *Communications biology*, 6(1), p. 451.

Varley, T.F., Pope, M., Maria Grazia, *et al.* (2023) 'Partial entropy decomposition reveals higher-order information structures in human brain activity', *Proceedings of the National Academy of Sciences of the United States of America*, 120(30), p. e2300888120.

Vinck, M. *et al.* (2023) 'Principles of large-scale neural interactions', *Neuron*, 111(7), pp. 987–1002.

Vossel, S., Geng, J.J. and Fink, G.R. (2014) 'Dorsal and ventral attention systems: distinct neural circuits but collaborative roles', *The Neuroscientist: a review journal bringing neurobiology, neurology and psychiatry*, 20(2), pp. 150–159.

Walton, M.E. *et al.* (2010) 'Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning', *Neuron*, 65(6), pp. 927–939.

Wang, R. *et al.* (2021) 'Segregation, integration, and balance of large-scale resting brain networks configure different cognitive abilities', *Proceedings of the National Academy of Sciences of the United States of America*, 118(23).

Watkins, C.J.C.H. and Dayan, P. (1992) 'Q-learning', *Machine learning*, 8(3), pp. 279–292.

White, J.K. *et al.* (2019) ‘A neural network for information seeking’, *Nature communications*, 10(1), p. 5168.

Wibral, M. *et al.* (2017) ‘Partial information decomposition as a unified approach to the specification of neural goal functions’, *Brain and cognition*, 112, pp. 25–38.

Williams, P.L. and Beer, R.D. (2010) ‘Nonnegative Decomposition of Multivariate Information’, *arXiv [cs.IT]*.

Wilson, R.C. *et al.* (2014) ‘Humans use directed and random exploration to solve the explore-exploit dilemma’, *Journal of experimental psychology. General*, 143(6), pp. 2074–2081.

Yu, A.J. and Dayan, P. (2005) ‘Uncertainty, neuromodulation, and attention’, *Neuron*, 46(4), pp. 681–692.

Yu, S. *et al.* (2011) ‘Higher-order interactions characterized in cortical activity’, *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 31(48), pp. 17514–17526.

Zajkowski, W.K., Kossut, M. and Wilson, R.C. (2017) ‘A causal role for right frontopolar cortex in directed, but not random, exploration’, *eLife*, 6.

Zamora-López, G. and Gilson, M. (2024) ‘An integrative dynamical perspective for graph theory and the analysis of complex networks’, *Chaos* , 34(4).