Behavioral/Cognitive

The Anterior Insula Engages in Feature- and Context-Level Predictive Coding Processes for Recognition Judgments

Cristiano Costa,[1] Cristina Scarpazza,[2,3] and Nicola Filippini[3]

[1]Padova Neuroscience Center, Università degli Studi di Padova, Padua 35131, Italy, [2]Dipartimento di Psicologia Generale, Università degli Studi di Padova, Padua 35131, Italy, and [3]IRCCS San Camillo Hospital, Venice 30126, Italy

Predictive coding mechanisms facilitate detection and perceptual recognition, thereby influencing recognition judgements, and, broadly, perceptual decision-making. The anterior insula (AI) has been shown to be involved in reaching a decision about discrimination and recognition, as well as to coordinate brain circuits related to reward-based learning. Yet, experimental studies in the context of recognition and decision-making, targeting this area and based on formal trial-by-trial predictive coding computational quantities, are sparse. The present study goes beyond previous investigations and provides a predictive coding computational account of the role of the AI in recognition-related decision-making, by leveraging Zaragoza-Jimenez et al. (2023) open fMRI dataset (17 female, 10 male participants) and computational modeling, characterized by a combination of view-independent familiarity learning and contextual learning. Using model-based fMRI, we show that, in the context a two-option forced–choice identity recognition task, the AI engages in feature-level (i.e., view-independent familiarity) updating and error signaling processes and context-level familiarity updating to reach a recognition judgment. Our findings highlight that an important functional property of the AI is to update the level of familiarity of a given stimulus while also adapting to task-relevant, contextual information. Ultimately, these expectations, combined with input visual signals through reciprocally interconnected feedback and feedforward processes, facilitate recognition judgments, thereby guiding perceptual decision-making.

Key words: anterior insula; decision-making; fMRI; perception; predictive coding; recognition

Significance Statement

Despite the renowned role of the anterior insula (AI) within the salience network and error-monitoring network, studies with a focus on this area and based on formal trial-by-trial predictive coding computational quantities are sparse. The present study provides a formal predictive coding computational account of the AI involvement in recognition-related decision–making. The present results demonstrate that AI activity reflects its engagement in encoding and updating the strength of an agent's belief in the statistical dependencies within the environment, thereby guiding perceptual decision-making. This underscores the pivotal role of the AI in integrating sensory information and mediating recognition-related decision–making processes. Overall, the findings highlight the AI's function in updating familiarity levels of stimuli and processing contextual information, ultimately facilitating recognition judgments.

Introduction

As perceptual agents, we navigate a reality constructed from a large amount of noisy and ambiguous sensory inputs (Clark, 2013; Walsh et al., 2020). Predictive coding provides a framework for understanding how we can accomplish such a task (Bastos et al., 2012; Keller and Mrsic-Flogel, 2018; Walsh et al., 2020). Drawing on past experiences, the brain proactively shapes perception by constructing feature-level predictions of expected inputs, which are iteratively updated and refined based on prediction errors (i.e., the residual mismatch between predictions and actual sensory inputs; Friston, 2008; Kok et al., 2014, 2017; Summerfield and de Lange, 2014; Gładziejewski, 2016). Over time, this process fosters learning by aligning predictions with statistical regularities in the environment while also accounting for contextual information through the integration of stimulus probability and contextually driven expectations with sensory inputs (Friston, 2005; Seth, 2013; Summerfield and de Lange, 2014; Barrett and Simmons, 2015; Sterzer et al., 2019).

Received May 9, 2024; revised Sept. 16, 2024; accepted Oct. 19, 2024.

Author contributions: C.C. designed research; C.C. and N.F. analyzed data; C.C., C.S., and N.F. wrote the paper.

The raw fMRI data and computational modeling used in this work were respectively acquired and developed by Zaragoza-Jimenez et al. (2023) and openly published in the Openneuro repository (https://doi.org/10.18112/openneuro.ds004529.v1.1.0). The contribution from N.F. to this work was supported by the Italian Ministry of Health (Ricerca Corrente). We thank the reviewers for their valuable insights.

The authors declare no competing financial interests.

Correspondence should be addressed to Cristiano Costa at cristiano.costa@phd.unipd.it.

https://doi.org/10.1523/JNEUROSCI.0872-24.2024

Ultimately, this mechanism facilitates adaptive behavior, detection, and perceptual recognition, thereby influencing recognition judgements and, broadly, perceptual decision-making (Ploran et al., 2007; Summerfield and de Lange, 2014; FitzGerald et al., 2015).

Among the currently identified brain regions involved in these processes, the anterior insula (AI) stands out as a pivotal hub (Bastin et al., 2016; Billeke et al., 2020; Loued-Khenissi et al., 2020), with consistent involvement across studies and stimulus features (Siman-Tov et al., 2019; Ficco et al., 2021; Corlett et al., 2022). Beyond its established role within the salience network (Menon and Uddin, 2010), the AI has been demonstrated to encode new and unanticipated stimuli, which are continuously tested through perceptual reconstruction and sampling, and to swiftly detect and convey error signals within the error-monitoring network (Downar et al., 2001; Crottaz-Herbette and Menon, 2006; Bastin et al., 2016; Billeke et al., 2020). Additionally, the AI has been shown to be involved in reaching a decision about discrimination and recognition (Ploran et al., 2007; Liu and Pleskac, 2011) and to coordinate brain circuits related to reward-based learning (Castelhano et al., 2014; Chand and Dhamala, 2016). Yet, experimental studies with a focus on the AI and based on formal trial-by-trial predictive coding computational quantities are sparse (Loued-Khenissi et al., 2020), and the nature of the predictive coding processes computed by this region warrants further investigation.

In this study, we go beyond previous investigations and provide a predictive coding computational account of the role of the AI in recognition-related decision–making. To achieve this, we leverage the publicly available dataset of functional magnetic resonance imaging (fMRI) data and predictive coding-consistent winning computational model parameters of Zaragoza-Jimenez et al. (2023; OSF repository, https://osf.io/tye24/).

Zaragoza-Jimenez et al. (2023) developed a series of computational models, both consistent and inconsistent with predictive coding principles, to model face identity learning. These models allowed for the computation of parameters such as learning rates and trial-by-trial estimates of hidden quantities, which were then related to brain activity in core face system regions, including the fusiform face area (FFA), occipital face area (OFA), and posterior superior temporal sulcus (pSTS; Haxby et al., 2000). The study aimed to explore how familiarity recognition is influenced by both stimulus familiarity and contextual information. In this context, stimulus familiarity was defined by the number of times a specific facial identity had been presented before a given trial, while contextual familiarity was defined by the average familiarity of stimuli presented immediately prior to the target stimulus, regardless of whether these preceding stimuli were the same as the target stimulus. Therefore, stimulus familiarity represents a stimulus-specific measure, whereas contextual familiarity reflects the broader environmental influence on recognition processes. The winning model, which best captured the learning process on a group level and was consistent with predictive coding principles, incorporated both view-independent processing, which assumes that representation of face identity emerges from the ability to predict different instances of the same face (Meaux et al., 2019) and the influence of contextual cues.

We hypothesize that the AI as well engages in computing feature-level (i.e., view-independent familiarity of a given stimulus) and context-level predictive coding quantities that reflect the gathering of information used to reach a recognition judgment (Li et al., 2009; Kahnt et al., 2011). Here, we assume that predictions and prediction errors are serially processed across hierarchical levels (Bastos et al., 2012; Weilnhammer et al., 2018). Initial perceptual information is analyzed in early sensory areas with respect to top–down predictive modulation. Prediction errors are then sent up the hierarchy to update higher-level representations, upon which top–down predictions are based (Bastos et al., 2012). As information progresses through the processing hierarchy, it becomes more complex (Perry and Fallah, 2014). Here, higher-order areas such as the AI likely encode high-level stimulus features and environmental dependencies, ultimately guiding recognition judgments (Loued-Khenissi et al., 2020).

Furthermore, the posterior insula (PI) was also examined to clarify the specific role of the anterior portion within the broader insular region in reflecting trial-by-trial predictive processes described in the present study.

Two specific sets of hypotheses were formulated:

1. The AI is hypothesized to update feature-level familiarity by maintaining predictive representations of upcoming sensory inputs (Allen et al., 2016; Fazeli and Büchel, 2018; Loued-Khenissi et al., 2020) and adjusting them when prediction errors arise (Wiech et al., 2010; Allen et al., 2016). Additionally, the AI is expected to process contextual information related to stimulus familiarity, as supported by findings across various domains (Wiech et al., 2010; Billeke et al., 2020). Overall, the AI integrates top–down predictive modulations with bottom–up signals, guiding decision-making processes based on stimulus familiarity (Chand and Dhamala, 2016; Loued-Khenissi et al., 2020). As such, its activity will reflect the following computational parameters derived from the winning model: view-independent familiarity $VI_{i,t}$, which represents the familiarity of a specific stimulus i at trial t, independent of the perspective from which the stimulus is presented, and is updated based on previous encounters with the stimulus; view-independent prediction error (δ), which represents the discrepancy between the maximum possible familiarity and the actual familiarity of the stimulus i at its previous presentation, signaling how much the current input deviates from what was expected and serving as an error signal in the updating process; contextual familiarity (C_t), which represents the updated estimate of how familiar the context is at trial t and is derived from the familiarity of stimuli presented in previous trials, adjusted by the contextual prediction error; and total familiarity (F_t), or the product of view-independent familiarity and contextual familiarity, providing an integrated measure of how familiar the current stimulus is, considering both its inherent familiarity and the influence of the surrounding context. We do not anticipate contextual prediction error (ε), which measures the discrepancy between the contextual familiarity at the last trial and the view-independent familiarity of the current stimulus, to modulate activity within the AI. This expectation is based on both the lack of supporting evidence in the literature for the AI performing this function and the mathematical formulation of ε. While ε is crucial for updating C_t, it does not itself directly drive the familiarity signal that influences recognition in the current trial. Instead, it modulates C_t, which, in turn, affects the familiarity signal (being part of F_t calculation). This nuanced role makes ε essential for the dynamic adjustment of context, but its influence on the familiarity signal is more indirect, occurring through its impact on C_t.

2. Since the PI functions as a gateway facilitating the transfer to higher-order processing without directly engaging in predictive coding processes (Wiech et al., 2010) and considering its recognized role in primarily processing interoceptive signals (Zhao et al., 2023), which are unlikely to occur in the present experimental context, activity within the PI is not expected to reflect any of the predictive coding processes described above.

Materials and Methods

In what follows, "Study participants," "Experimental paradigm," "Computational modeling," and "fMRI data acquisition" sections are directly adapted from Zaragoza-Jimenez et al. (2023), and only pertinent information to the present study will be reported. For a thorough description and results of the power analysis, a comprehensive explanation of the complete experimental paradigm, and the implementation and detailed mathematical description of the computational modeling, please consult Zaragoza-Jimenez et al. (2023). Original code implementations, raw data, and analyses can be found in the associated public repository DOI (https://doi.org/10.18112/openneuro.ds004529.v1.1.0).

The section "fMRI data analysis" describes the novel contribution of the present work.

Study participants

Thirty-four healthy participants (21 females, 13 males; mean age 23.8 ± 2.6 years) met the inclusion criteria. Following the protocol established by Zaragoza-Jimenez et al. (2023), two participants were excluded due to poor MRI data quality, three due to extremely high beta values for view-independent familiarity ($VI_{i,t}$), and one for failing to achieve a minimum correct face-recognition rate of 60% averaged over all stimuli by the end of the experiment. Additionally, upon inspection, we excluded an additional participant from the analysis due to out-of-range parameters: total familiarity (F_t) exceeded the group mean by 3.2 standard deviations, view-independent familiarity ($VI_{i,t}$) exceeded the group mean by 2.4 standard deviations, and contextual familiarity (C_t) exceeded the group mean by 2.1 standard deviations. Consequently, the final sample for the present study comprised 27 participants (17 females, 10 males; mean age 23.9 ± 2.5 years).

Experimental paradigm

The faces adopted in the task were sourced from the "300 random faces" dataset (Todorov et al., 2008), accessible within the database of the Social Perception Lab at Princeton University (http://tlab.princeton.edu/databases/secretdatabaseportal/). Participants underwent a face-recognition task where they were repeatedly shown computer-generated faces in varying viewpoints (front-facing or at a 30° tilt from either side). Stimuli were presented centrally against a uniform gray background. After each presentation, participants were required to indicate whether they had seen that person in a two-option forced–choice task, regardless of viewpoint. Twenty-four different face identities were presented, initially unknown to participants. Fifteen faces appeared multiple times from various viewpoints (front-facing or at a 30° tilt from either side), totaling 12 presentations per face. Conversely, nine face identities were shown only once, with three presented front-facing and three from each side angle. In total, 189 face stimuli were presented.

Face stimuli were ordered so that the identities of the faces became gradually more familiar to the participants throughout the experiment. The stimulus order over time was based on the ground truth of the task, which comprised stimulus familiarity (indicating how many times a specific face stimulus had been presented before) and contextual familiarity (representing the average familiarity of previously presented faces). Both parameters were systematically varied across trials. Identity sequence and viewpoint presentation were pseudorandomized, preventing participants from anticipating upcoming stimuli beyond contextual familiarity predictions. Contextual familiarity was systematically controlled by manipulating the average frequency a face identity was seen before, ensuring randomization across viewpoints and identities. Additionally, faces shown only once were deferred to later stages of the experiment to prevent accuracy improvement through increased "yes" responses.

Computational modeling

The models developed by Zaragoza-Jimenez et al. (2023) are characterized by different combinations of three main computational processes: view-dependent familiarity learning, view-independent familiarity learning, and contextual learning. View-dependent processing involves encoding face-related visual information directly as it reaches the retina, without significant higher-order abstraction. This means that when the same face is observed from different angles, each angle induces a distinct stimulus representation, rather than being perceived as different instances of the same stimulus. In contrast, view-independent processing, implemented in the winning model, relies more on abstract higher-level inference in visual perception. In this case, it is hypothesized that the brain utilizes a single abstract representation for face identification. Seeing one particular instance of a stimulus triggers the retrieval of related information, generating sensory predictions about other features associated with it. Thus, the representation of face identity emerges from the ability to predict different instances of the same face (Meaux et al., 2019).

The data were fitted subject-wise using a maximum-likelihood approach. The joint-likelihood of the observations obtained from each participant under each model was evaluated using the sum of the log-likelihoods for the trial-wise model predictions. Parameter optimization for maximizing the joint probability of the observations under each model was done using the L-BFGS-B algorithm with a finite-differences numerical gradient approximation, which included Basin-hopping for improved global optimization. All parameters were implemented to vary in steps of 0.01.

Each model comprised a hidden-variable model and a response model. While the response model remained consistent across all models, the architecture of the hidden-variable model varied depending on the computational model. The hidden-variable model's function was to update and carry the trial-by-trial estimates of the hidden variables and convert them into the total familiarity of a stimulus at a given trial. Meanwhile, the response model converted the trial-by-trial total stimulus familiarity predictions output by the model into probabilities for the two possible actions of the task ("Yes" or "No") using the Luce-choice rule (Luce, 1977; Luce et al., 2008). The model-predicted probability of answering "Yes" at a given trial was calculated as follows:

$$p_{\mathrm{Yes}(t)} = \frac{1}{1 + e^{(-\beta F_t)}}, \tag{1}$$

where β represents the stochastic parameter indicating the sensitivity of the answer-behavior to the total familiarity of the stimulus at time t (F_t). As implied by the binary answer format, the probability of answering "No" is given as follows:

$$p_{\mathrm{No}(t)} = 1 - p_{\mathrm{Yes}(t)}. \tag{2}$$

The model architecture requires initialization values to start the hidden-variable updating processes. These initialization values can be understood as the prior information available to an individual when starting the cognitive task. Explicit information was provided for the view-independent familiarity where the initial hidden-variable state $V_{i,0}$ for all stimuli i was defined as follows:

$$V_{i,0} = \mathrm{FPrate} * \lambda. \tag{3}$$

Here, FPrate represents the percentage of false-positive responses for those stimuli presented for the first time, regardless of the view angle, indicating incorrect recognition of a stimulus as familiar during its initial presentation. λ denotes the individual parameter reflecting the maximum view-independent familiarity achievable for any stimulus in this participant.

The initial contextual familiarity was set to 0, reflecting the total unfamiliarity of each participant with the stimulus material at the beginning of the task.

In the winning model, namely, the view-independent and context-dependent model, which demonstrated the highest approximate Bayes factor ($ABF_{10} = 9999.50$) during model selection, the overall familiarity of a stimulus is determined by both its view-independent familiarity (representing perspective-invariant face identity learning) and contextual familiarity (reflecting stimulus familiarity across preceding trials). The total familiarity (F_t) in this model is calculated as the product of these two components:

$$F_t = \mathrm{VI}_{i,t} * C_t, \tag{4}$$

where $\mathrm{VI}_{i,t}$ is the view-independent familiarity of stimulus i at trial t. $\mathrm{VI}_{i,t}$ is hypothesized to be computed after the presentation of each stimulus and is subsequently utilized in calculating contextual familiarity and, ultimately, total familiarity. The calculation of $\mathrm{VI}_{i,t}$ is given by the following formula:

$$\mathrm{VI}_{i,t} = \mathrm{VI}_{i,n-1} + \alpha\delta, \tag{5}$$

where $\mathrm{VI}_{i,n-1}$ denotes the view-independent familiarity of a stimulus i at its previous presentation, α denotes the view-independent learning rate, and δ denotes the view-independent prediction error. δ is calculated as follows:

$$\delta = \lambda - \mathrm{VI}_{i,n-1}, \tag{6}$$

where λ is the individual maximum view-independent familiarity. Contextual familiarity at trial t is calculated as follows:

$$C_t = C_{t-1} - \sigma\varepsilon, \tag{7}$$

where C_{t-1} represents the contextual familiarity at the previous trial, σ denotes the contextual learning rate, and ε signifies the contextual prediction error, computed as follows:

$$\varepsilon = C_{t-1} - \mathrm{VI}_{i,t}, \tag{8}$$

Once again C_{t-1} represents the contextual familiarity at the last trial, and $\mathrm{VI}_{i,t}$ denotes the view-independent familiarity of stimulus i at the current trial, calculated using the formula in Equation 5.

fMRI data acquisition

MRI data were acquired using a Siemens 3 Tesla Tim Trio MRI scanner (Siemens Medical Systems) equipped with a 32-channel head matrix receive coil at the Core Facility Brain Imaging, University of Marburg. Functional images were acquired using a T2*-weighted gradient–echo echoplanar imaging sequence sensitive to the blood-oxygenation-level-dependent (BOLD) contrast. The imaging parameters were as follows: repetition time (TR), 1,530 ms; echo time (TE), 30 ms; voxel size, $3 \times 3 \times 3$ mm^3; and flip angle, 55°. Slices were acquired in an ascending order parallel to the intercommissural plane (anterior to posterior commissure) using a multiband acceleration factor of 2. The first four images for the face identity learning task were discarded to eliminate the effects of T1 stabilization. Prior to functional imaging, a high-resolution T1–weighted anatomical scan covering the entire brain was acquired using a magnetization-prepared rapid gradient-echo (3D) sequence in the sagittal plane. The parameters for the anatomical scan were as follows: TR, 1,900 ms; TE, 2.26 ms; voxel size, $1 \times 1 \times 1$ mm^3; and flip angle, 9°.

fMRI data analysis

Image processing. First-level fMRI analysis was carried out using fMRI Expert Analysis Tool (FEAT) v. 6.00 (Woolrich et al., 2001). Preprocessing consisted of head motion correction, brain extraction, spatial smoothing using a Gaussian kernel of full-width at half-maximum of 5 mm, and high-pass temporal filtering equivalent to 128 s. Time-series statistical analysis was carried out with local autocorrelation correction. A boxcar convolved with a gamma hemodynamic response function and its temporal derivative was used to model the data. Five separate first-level analyses were conducted for each experimental participant. In each analysis, the first explanatory variable (EV) included the onset and duration of the face stimuli, as well as the specific trial-by-trial computational parameter under investigation (i.e., $\mathrm{VI}_{i,t}$, δ, C_t, F_t, and ε). This EV represented the main contrast of interest, with the specific aim of investigating positive BOLD activation associated with the weight attributed to each trial. Additionally, we included a second EV as a nuisance parameter to account for individual variability in the mean BOLD signal. This second EV was orthogonalized with respect to the contrast of interest.

FMRI volumes were registered to the individual's structural scan and standard space images using both linear (FLIRT) and nonlinear (FNIRT) registration tools and then optimized using boundary-based–registration approach (Greve and Fischl, 2009). The derived transformations into standard space were applied to images of contrasts of interest and their variances. Preprocessing results of each of the study participants were visually inspected by a trained neuroscientist (NF) to ensure registration accuracy.

Higher-level (group-level) whole–brain analysis was carried out using FMRIB's Local Analysis of Mixed Effects (Woolrich et al., 2004). For each of the five computational parameters investigated, we tested for group averages for each of the contrasts of interest. The Z statistic images were thresholded with clusters defined by $Z > 3.1$, and a family-wise error–corrected cluster significance threshold of $p < 0.05$ was applied to the suprathreshold clusters.

Region of interest (ROI) analysis. A total of five pairs of anatomically identical ROIs were created, with five located in the left hemisphere and five in the right hemisphere. ROIs were defined as spheres with a 5 mm radius centered around their respective centroids. Centroids were identified from the Multilevel Human Atlas (https://www.ebrains.eu/brain-atlases/reference-atlases/human-brain/), using the ICBM 2009c Nonlinear Symmetric template and Julich-Brain parcellation (v3.0.3; Amunts et al., 2020). ROI pairs included the following brain areas: AI (Area Id7), PI (Area Id2), FFA (mid-anterior fusiform gyrus), OFA (inferior occipital gyrus), and pSTS. The BOLD percentage signal change (to carry out statistical analyses) from the five first-level fMRI data of each participant was extracted from each selected ROI using Featquery (part of FSL).

The three core face system ROI pairs (i.e., FFA, OFA, and pSTS) previously analyzed by Zaragoza-Jimenez et al. (2023) were included in the present analysis to assess the robustness of effects considering the methodological and analytical differences with respect to Zaragoza-Jimenez et al. (2023) investigation. Unlike Zaragoza-Jimenez et al. (2023), which used functional ROIs and conducted analyses in SPM12 (www.fil.ion.ucl.ac.uk/spm), our study utilized anatomically defined ROIs. For the first-level analysis, the authors' general linear model included the following conditions: the first regressor modeled the onset the stimuli, including only events in which a response was given; the second regressor modeled the onset of the trigger cue events (question: "Do you recognize the person?"), again including only events in which a response was given; the third regressor modeled both the aforementioned events on missed trials; the fourth regressor described the effect of specific computational parameters (for detailed methodological specifics, refer to Zaragoza-Jimenez et al., 2023). Therefore, the present results for the core face system ROIs should not be considered a direct replication attempt of Zaragoza-Jimenez et al. (2023) results.

Primary analysis: modulation of the BOLD percentage signal change. The following analysis was implemented in the statistical environment R (version 2023.06.0+421; R Core Team, 2023) and used the following packages: dplyr (Wickham et al., 2023), tidyr (Wickham et al., 2024), and broom (Robinson et al., 2023).

Firstly, group-level mean BOLD percentage signal change values were computed for each ROI and computational parameter combination to provide a comprehensive understanding of the data distribution within each region. One-sample t tests were executed to assess whether the mean BOLD percentage signal change, at the group level, significantly differed from zero within each ROI for each computational

parameter. Cohen's *d* was calculated to estimate the effect size, representing the magnitude of the observed effect relative to the variability in the data. The significance level was set at $\alpha = 0.0001$ for all statistical tests, controlling the family-wise error rate through Bonferroni's correction to mitigate the risk of Type 1 errors.

Secondly, pairwise comparisons were then performed between the BOLD percentage signal change associated with each computational parameter within each ROI to determine whether any specific computational parameter showed significantly different associations with the BOLD signal compared with the other computational parameters. For each ROI, the BOLD percentage signal changes associated with each pair of computational parameters were compared using paired *t* tests. Effect sizes were computed using Cohen's *d*, where the pooled standard deviation was used to account for variability across the computational parameters. The significance level was again set at $\alpha = 0.0001$ for the pairwise comparisons, with Bonferroni's correction applied to adjust for multiple comparisons.

Post hoc analysis: investigating laterality differences. Post hoc tests and dataframe manipulation were carried out within the statistical environment R by means of the following packages: lme4 (Bates et al., 2015), multcomp (Hothorn et al., 2008), broom (Robinson et al., 2023), and dplyr (Wickham et al., 2023).

Following the primary statistical analyses, post hoc tests were conducted iteratively for each left–right ROI pair within each computational parameter to investigate potential laterality differences. Here, laterality differences refer to differences in BOLD percentage signal change between the same ROIs located on the left and right sides of the brain. The procedure involved subsetting the data for the specific computational parameter and ROI pair, fitting a linear mixed-effect model, and performing post hoc tests using Tukey's method for multiple comparisons. The significance level was set at $\alpha = 0.0001$ for all statistical tests.

Results

fMRI data analysis

Face identity learning task

Whole-brain group analysis of task-based fMRI data for the five computational parameters investigated [i.e., view-independent familiarity ($VI_{i,t}$), view-independent prediction error (δ), contextual familiarity (C_t), total familiarity (F_t), and contextual prediction error (ε)] revealed that four out of five of the computational parameters tested had specific brain regions involved during the task, with a varying degree of BOLD activation depending on the specific computational parameter. Whereas ε was not associated with any brain activation, brain regions commonly associated with the other four computational parameters included the following brain areas: the occipital pole extending to the temporal fusiform cortex bilaterally, the superior parietal lobule bilaterally, the precentral gyrus bilaterally, the insular cortex bilaterally, the middle frontal gyrus bilaterally, the putamen and caudate bilaterally, and the paracingulate gyrus (Fig. 1). All these regions survived the statistical threshold.

Noteworthy, as reported in Zaragoza-Jimenez et al. (2023), the computational parameters partly covary across trials. Specifically, the authors computed the group-level mean correlation calculated over the trial-by-trial parameters, which showed that F_t highly correlates with C_t ($r = 0.976$) and $VI_{i,t}$ ($r = 0.956$), as expected since $F_t = VI_{i,t} * C_t$. Additionally, C_t highly correlates with $VI_{i,t}$ ($r = 0.94$). While these intercorrelations are not surprising, they highlight the challenge of attributing significance to a single parameter without accounting for the contributions of the other correlated parameters.

Primary analysis: modulation of percentage signal change

The primary analysis aimed at investigating whether the BOLD percentage signal change within our ROIs was influenced by parameters derived from the winning computational model, i.e., the view-independent and context-dependent model. Our hypotheses were as follows:

1. Activity within the AI would reflect $VI_{i,t}$, δ, C_t, and F_t, resulting in an increase in mean BOLD percentage signal change. We did not expect ε to influence activity within the AI.
2. Activity within the PI would not reflect any computational parameters.

A summary of statistics (i.e., *t* statistic, effect size, mean percentage signal change, and *p* value) can be found in Table 1. BOLD percentage signal changes for each computational parameter and ROI are plotted in Figure 2. Our results support Hypothesis 1. The BOLD percentage signal change in the AI, both left and right, was significantly greater than zero as a function of $VI_{i,t}$ (left, $t_{(26)} = 12.695$; $p < 0.0001$; Cohen's $d = 2.443$; right, $t_{(26)} = 11.976$; $p < 0.0001$; Cohen's $d = 2.305$), δ (left, $t_{(26)} = 12.253$; $p < 0.0001$; Cohen's $d = 2.358$; right, $t_{(26)} = 10.504$; $p < 0.0001$; Cohen's $d = 2.022$); C_t (left, $t_{(26)} = 8.645$; $p < 0.0001$; Cohen's $d = 1.664$; right, $t_{(26)} = 8.865$; $p < 0.0001$; Cohen's $d = 1.706$), and F_t (left, $t_{(26)} = 10.392$; $p < 0.0001$; Cohen's $d = 2$; right, $t_{(26)} = 9.773$; $p < 0.0001$; Cohen's $d = 1.881$), but not of ε (left, $t_{(26)} = 0.091$; $p < 0.928$; Cohen's $d = 0.018$; right, $t_{(26)} = 0.672$; $p < 0.507$; Cohen's $d = 0.129$).

Activity within the PI, both in the left and right hemispheres (Extended Data Table 1-1), showed BOLD percentage signal changes significantly lower than zero. Specifically, left PI BOLD percentage signal change significantly decreased below zero as a function of $VI_{i,t}$ ($t_{(26)} = -4.880$; $p < 0.0001$; Cohen's $d = 0.939$), while right PI BOLD percentage signal change decreased below zero as a function of $VI_{i,t}$ ($t_{(26)} = -6.796$; $p < 0.0001$; Cohen's $d = 1.308$), δ ($t_{(26)} = -4.932$; $p < 0.0001$; Cohen's $d = 0.949$), C_t ($t_{(26)} = -6.198$; $p < 0.0001$; Cohen's $d = 1.193$), and F_t ($t_{(26)} = -5.736$; $p < 0.0001$; Cohen's $d = 1.104$). Thus, Hypothesis 2 is rejected.

Regarding the core face system ROIs (i.e., FFA, OFA, pSTS, Extended Data Table 1-1), the BOLD percentage signal change in both left and right FFA was significantly greater than zero as a function of $VI_{i,t}$ (left, $t_{(26)} = 10.853$; $p < 0.0001$; Cohen's $d = 2.089$; right, $t_{(26)} = 8.845$; $p < 0.0001$; Cohen's $d = 1.702$), δ (left, $t_{(26)} = 9.273$; $p < 0.0001$; Cohen's $d = 1.785$; right, $t_{(26)} = 6.224$; $p < 0.0001$; Cohen's $d = 1.198$), C_t (left, $t_{(26)} = 9.130$; $p < 0.0001$; Cohen's $d = 1.757$; right, $t_{(26)} = 8.751$; $p < 0.0001$; Cohen's $d = 1.684$), and F_t (left, $t_{(26)} = 9.434$; $p < 0.0001$; Cohen's $d = 1.816$; right, $t_{(26)} = 8.677$; $p < 0.0001$; Cohen's $d = 1.670$). Regarding the OFA, BOLD percentage signal change was significantly greater than zero in the right region, as a function of $VI_{i,t}$ ($t_{(26)} = 9.258$; $p < 0.0001$; Cohen's $d = 1.782$), δ ($t_{(26)} = 10.140$; $p < 0.0001$; Cohen's $d = 1.951$), C_t ($t_{(26)} = 8.184$; $p < 0.0001$; Cohen's $d = 1.575$), and F_t, ($t_{(26)} = 7.882$; $p < 0.0001$; Cohen's $d = 1.517$). On the other hand, none of the computational parameters modulated activity within the pSTS.

Table 2 and Extended Data Table 2-1 report the statistical features (i.e., *t* statistic, effect size, and *p* value) of the pairwise comparisons between the BOLD percentage signal changes associated with each pair of computational parameters within each ROI. The BOLD percentage signal change associated with ε was significantly weaker ($p < 0.0001$) than the BOLD percentage signal change associated with $VI_{i,t}$ ($t = -7.175$; Cohen's $d = 2.141$), δ ($t = -7.496$; Cohen's $d = 2.123$), C_t ($t = -5.020$; Cohen's $d = 1.678$), and F_t ($t = -5.541$; Cohen's $d = 1.811$) within the left AI; $VI_{i,t}$ ($t = -6.349$; Cohen's $d = 1.854$) and δ ($t = -6.545$; Cohen's $d = 1.814$) within the right

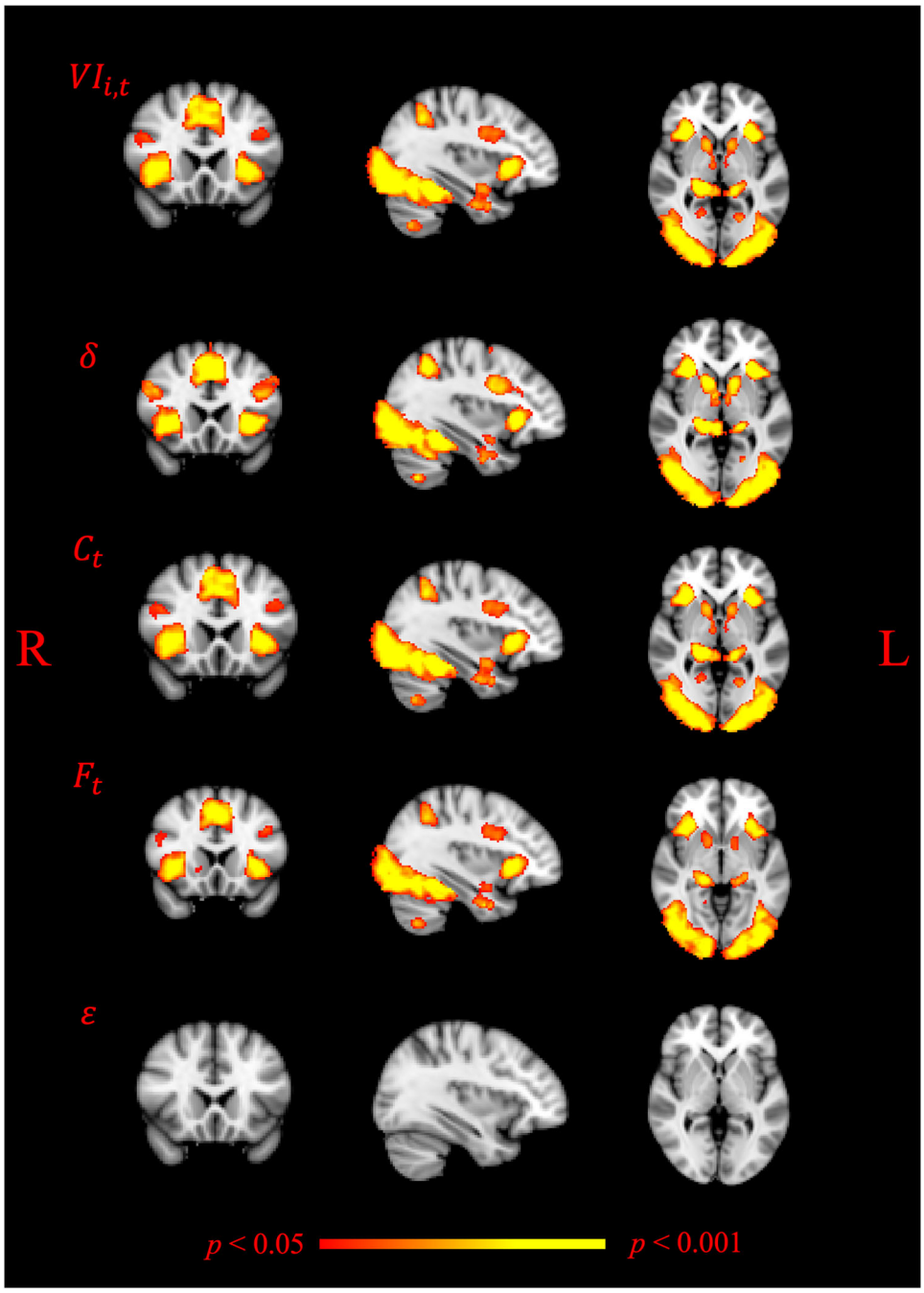

Figure 1. Whole-brain results of group analysis for the five computational parameters investigated. $VI_{i,t}$, view-independent familiarity; δ, view-independent prediction error; C_t, contextual familiarity; F_t, total familiarity; ε, contextual prediction error. All images displayed here report results with a threshold of $Z > 3.1$, with cluster-based significance corrected for multiple comparison using family-wise error correction at $p < 0.05$ (red-to-yellow color bar). R, right hemisphere; L, left hemisphere.

Table 1. One-sample *t* tests results for the AI

ROI	Parameter	*t* statistic	Degrees of freedom	Effect size	Mean % signal change	*p* value	Corrected *p* value
L AI	$VI_{i,t}$	12.695	26	2.443	0.289	<0.0001	<0.0001
L AI	δ	12.253	26	2.358	0.291	<0.0001	<0.0001
L AI	C_t	8.645	26	1.664	0.255	<0.0001	<0.0001
L AI	F_t	10.392	26	2.000	0.250	<0.0001	<0.0001
L AI	ε	0.091	26	0.018	0.003	0.928	1
R AI	$VI_{i,t}$	11.976	26	2.305	0.312	<0.0001	<0.0001
R AI	δ	10.504	26	2.022	0.328	<0.0001	<0.0001
R AI	C_t	8.865	26	1.706	0.284	<0.0001	<0.0001
R AI	F_t	9.773	26	1.881	0.268	<0.0001	<0.0001
R AI	ε	0.672	26	0.129	0.023	0.5073	1

Significance level: $\alpha = 0.0001$ family-wise error corrected. L AI, left anterior insula; R AI, right anterior insula; $VI_{i,t}$, view-independent familiarity; δ, view-independent prediction error; C_t, contextual familiarity; F_t, total familiarity; ε, contextual prediction error. See Extended Data Table 1-1 for all other ROIs results.

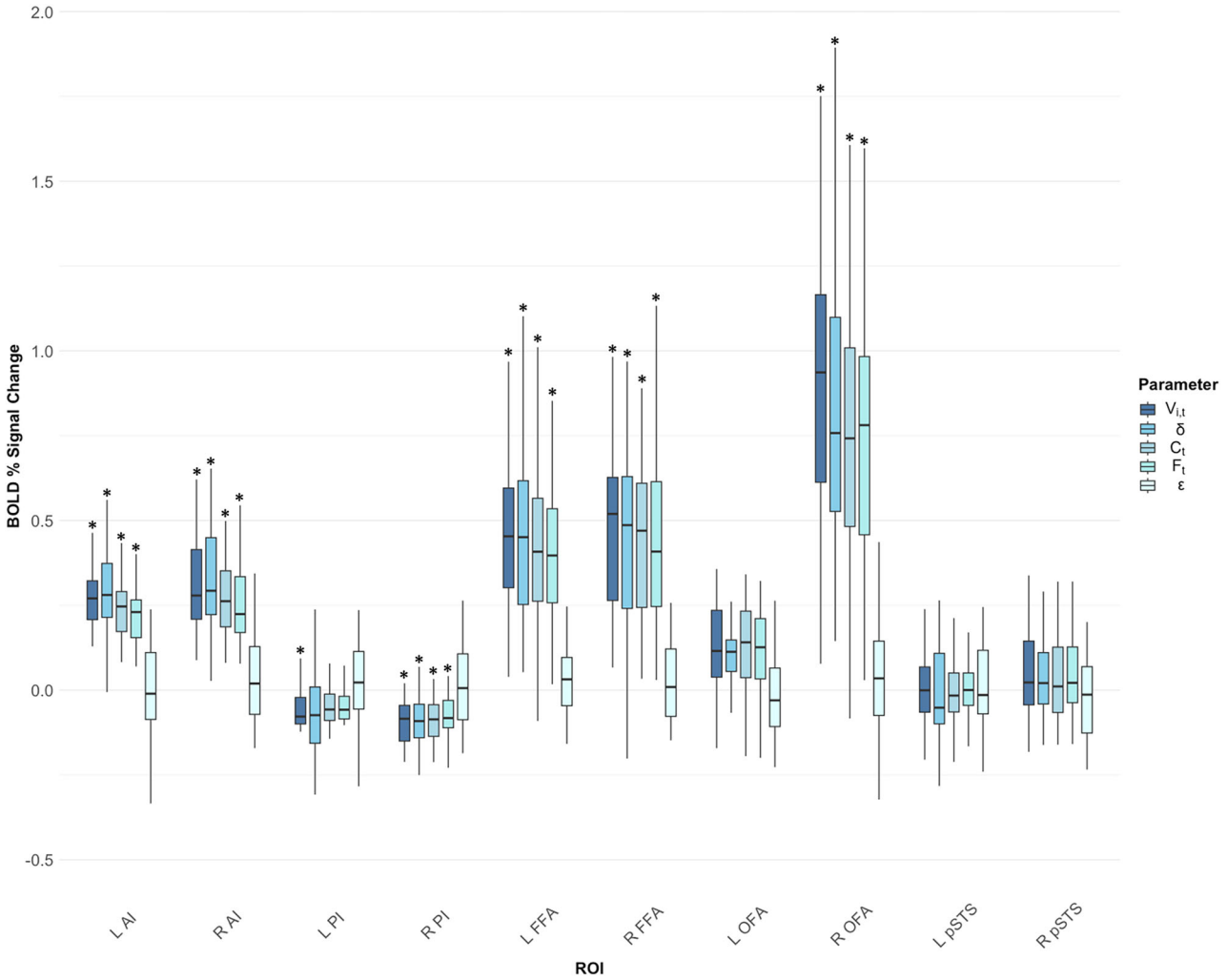

Figure 2. Box plots of BOLD percentage signal changes associated with each computational parameter within each ROI. R, right hemisphere; L, left hemisphere; AI, anterior insula; PI, posterior insula; FFA, fusiform face area; OFA, occipital face area; pSTS, posterior superior temporal sulcus; $VI_{i,t}$, view-independent familiarity; δ, view-independent prediction error; C_t, contextual familiarity; F_t, total familiarity; ε, contextual prediction error. Asterisks (*) indicate statistically significant ($p < 0.0001$) one-sample t test results against zero.

AI; δ ($t = -5.1$; Cohen's $d = 0.772$) within the left PI; $VI_{i,t}$ ($t = -7.791$; Cohen's $d = 2.353$), δ ($t = -7.332$; Cohen's $d = 2.091$), C_t ($t = -6.292$; Cohen's $d = 2.060$), and F_t ($t = -6.253$; Cohen's $d = 2.066$) within the left FFA; $VI_{i,t}$ ($t = -7.791$; Cohen's $d = 2.353$), δ ($t = -7.332$; Cohen's $d = 2.091$), C_t ($t = -6.292$; Cohen's $d = 2.060$), and F_t ($t = -6.253$; Cohen's $d = 2.066$) within the right FFA; and $VI_{i,t}$ ($t = -6.617$; Cohen's $d = 2.060$), δ ($t = -6.939$; Cohen's $d = 2.115$), C_t ($t = -5.532$; Cohen's $d = 1.852$), and F_t ($t = -5.299$; Cohen's $d = 1.777$) within the right OFA.

Post hoc analysis: investigating laterality differences

Post hoc tests were conducted to investigate differences in BOLD percentage signal change for each left–right ROI pair within each computational parameter using linear mixed-effect models.

Test statistics (i.e., estimate, standard error, Z-score, p value) are reported in Table 3. Laterality effects were observed in the OFA, where the BOLD percentage signal change was significantly higher in the right region compared with the left one for $VI_{i,t}$ [estimate (SD) = 0.778 (0.103); $p < 0.0001$; Z-score = 7.533], δ [estimate (SD) = 0.724 (0.092); $p < 0.0001$; Z-score = 7.858], C_t [estimate (SD) = 0.713 (0.107); $p < 0.0001$; Z-score = 6.666], and F_t [estimate (SD) = 0.680 (0.102); $p < 0.0001$; Z-score = 6.638].

Discussion

The AI role in recognition judgments

In the present study, we provided a formal predictive coding computational account of the role of the AI within the context of recognition-related decision–making. We hypothesized that the AI engages in feature-level (i.e., view-independent familiarity) updating and error signaling processes and context-level familiarity updating to reach a recognition judgment, thereby guiding perceptual decision-making (Kahnt et al., 2011; Jia et al., 2018). Our results support this hypothesis. Indeed, we observed a statistically significant increase in BOLD percentage signal change in both left and right AI as a function of view-independent familiarity ($VI_{i,t}$), view-independent prediction error (δ), contextual familiarity (C_t), and total familiarity (F_t). Additionally, the association between contextual prediction error (ε) and BOLD percentage signal change in this area was generally significantly weaker than the association with the other

Table 2. Paired *t* tests results between the BOLD percentage signal changes associated with each pair of computational parameters within each ROI

ROI	Parameter contrast	*t* statistic	Effect size	Corrected *p* value
L AI	$VI_{i,t}$ vs δ	−0.155	0.023	1
L AI	$VI_{i,t}$ vs C_t	2.243	0.248	1
L AI	$VI_{i,t}$ vs F_t	3.977	0.316	0.16
L AI	$VI_{i,t}$ vs ε	7.175	2.141	<0.0001
L AI	δ vs C_t	1.558	0.263	1
L AI	δ vs F_t	1.989	0.331	1
L AI	δ vs ε	7.496	2.123	<0.0001
L AI	C_t vs F_t	0.378	0.033	1
L AI	C_t vs ε	5.020	1.678	<0.0001
L AI	F_t vs ε	5.541	1.811	<0.0001
R AI	$VI_{i,t}$ vs δ	−0.855	0.109	1
R AI	$VI_{i,t}$ vs C_t	1.968	0.187	1
R AI	$VI_{i,t}$ vs F_t	4.015	0.313	0.16
R AI	$VI_{i,t}$ vs ε	6.349	1.854	<0.0001
R AI	δ vs C_t	1.926	0.271	1
R AI	δ vs F_t	2.674	0.391	1
R AI	δ vs ε	6.545	1.814	<0.0001
R AI	C_t vs F_t	1.827	0.098	1
R AI	C_t vs ε	4.648	1.532	0.032
R AI	F_t vs ε	4.753	1.543	0.032

Significance level: $\alpha = 0.0001$ corrected for multiple comparisons. L AI, left anterior insula; R AI, right anterior insula; $VI_{i,t}$, view-independent familiarity; δ, view-independent prediction error; C_t, contextual familiarity; F_t, total familiarity; ε, contextual prediction error. See Extended Data Table 2-1 for all other pairwise comparisons results.

Table 3. Linear mixed-effect model results

Contrast	Parameter	Estimate	Standard error	*Z*-score	Corrected *p* value
R AI vs L AI	$VI_{i,t}$	0.023	0.026	0.889	0.374
R PI vs L PI	$VI_{i,t}$	−0.033	0.013	−2.479	0.013
R FFA vs L FFA	$VI_{i,t}$	0.037	0.070	0.529	0.597
R pSTS vs L pSTS	$VI_{i,t}$	0.059	0.040	1.490	0.136
R OFA vs L OFA	$VI_{i,t}$	0.778	0.103	7.533	<0.0001
R AI vs L AI	δ	0.037	0.025	1.498	0.134
R PI vs L PI	δ	−0.013	0.020	−0.627	0.531
R FFA vs L FFA	δ	0.026	0.081	0.316	0.752
R pSTS vs L pSTS	δ	0.036	0.040	0.899	0.369
R OFA vs L OFA	δ	0.724	0.092	7.858	<0.0001
R AI vs L AI	C_t	0.029	0.025	1.159	0.246
R PI vs L PI	C_t	−0.035	0.015	−2.375	0.018
R FFA vs L FFA	C_t	0.029	0.065	0.444	0.657
R pSTS vs L pSTS	C_t	0.059	0.036	1.634	0.102
R OFA vs L OFA	C_t	0.713	0.107	6.666	<0.0001
R AI vs L AI	F_t	0.018	0.024	0.768	0.442
R PI vs L PI	F_t	−0.033	0.013	−2.546	0.011
R FFA vs L FFA	F_t	0.028	0.064	0.430	0.667
R pSTS vs L pSTS	F_t	0.056	0.034	1.656	0.098
R OFA vs L OFA	F_t	0.680	0.102	6.638	<0.0001
R AI vs L AI	ε	0.020	0.015	1.338	0.181
R PI vs L PI	ε	−0.004	0.018	−0.199	0.842
R FFA vs L FFA	ε	−0.040	0.048	−0.838	0.402
R pSTS vs L pSTS	ε	−0.031	0.028	−1.120	0.263
R OFA vs L OFA	ε	−0.017	0.076	−0.216	0.829

Significance level: $\alpha = 0.0001$ corrected for multiple comparisons. R, right hemisphere; L, left hemisphere; AI, anterior insula; PI, posterior insula; FFA, fusiform face area; OFA, occipital face area; pSTS, posterior superior temporal sulcus; $VI_{i,t}$, view-independent familiarity; δ, view-independent prediction error; C_t, contextual familiarity; F_t, total familiarity; ε, contextual prediction error.

computational parameters. These results indicate that AI activity reflects its role in encoding the strength of an agent's belief (and trial-by-trial beliefs updating and refinement processes) in the statistical dependencies within their environment and its involvement in aspects of learning and decision (Sayal et al., 2020; Verdade et al., 2022).

The AI has garnered attention as a pivotal hub in coordinating the integration of sensory information and mediating cognitive processing (Menon and Uddin, 2010; Seeley, 2019; Verdade et al., 2022). This involvement favors a discussion about its crucial role in processes such as recognition judgments and decision-making (Sterzer and Kleinschmidt, 2010; Liu and Pleskac, 2011; Sayal et al., 2020; Verdade et al., 2022). From the perspective of predictive coding models, perceptual decision-making is known to be a complex function mediated by a network consisting of dissociable but interacting processes (Mostert et al., 2015). In this framework, recognition judgments (i.e., inference) rely on predictions aimed at explaining away the discrepancy between predicted and incoming sensory information (Hesselmann et al., 2010). Accordingly, precise prediction errors, accumulated to optimize high-level representations and subsequent top–down predictions of sensory inputs, will bias toward correct inference (Friston, 2005). Our findings suggest that the AI contributes to these processes by likely maintaining multiple predictive representations of upcoming stimuli, updating them based on prediction errors generated whenever bottom–up sensory input deviates from prior expectations (Allen et al., 2016; Fazeli and Büchel, 2018). In particular, the view-independent and context-dependent predictive coding model implemented in the present work (and that captured the learning process best on the group level; see Zaragoza-Jimenez et al., 2023) assumes that, with repeated exposure, stimuli become more familiar, necessitating less updating ($VI_{i,t}$) for subsequent recognition. Conversely, the model predicts higher prediction errors (δ) evoked by unfamiliar stimulus identities than contextually probable ones. Our results are consistent with previous studies that investigated these processes in different domains (Ishai et al., 2000; Allen et al., 2016; Loued-Khenissi et al., 2020). Predictions are then presumably relayed via feedback connections to sensory areas (Weilnhammer et al., 2018). Specifically, feature-level predictions enhance early sensory representations and facilitate inference for contextually probable stimuli while suppressing alternatives (Dunovan et al., 2014; Watanabe and Sasaki, 2015), thereby guiding accurate decision-making about recognition.

Our findings also indicate that the AI processes contextual information (C_t) regarding stimulus familiarity, consistent with previous findings within the domain of pain perception (Wiech et al., 2010). An important function of the AI may therefore be to also process predictive information about task-relevant, contextual information (Wiech et al., 2010). Moreover, activity in the AI reflected the familiarity of each stimulus (F_t) as a function of both $VI_{i,t}$ and C_t, resulting in activity that varies according to the likelihood that a participant will perceive a stimulus as recognized (Apps and Tsakiris, 2013). This underscores its overarching functional property of signaling the salience of information that guides decisions (Wiech et al., 2010; Chand and Dhamala, 2016; Loued-Khenissi et al., 2020).

The present results show that the contextual prediction error (ε) was not reflected in the AI activity, supporting our hypothesis. This hypothesis was based on the lack of supporting evidence in the available literature for an AI's role in computing this type of predictive coding quantity and was also grounded in mathematical considerations. While ε is used to update C_t, which in turn affects the familiarity signal, ε does not itself directly drive the familiarity signal that influences recognition in the current trial. Instead, ε plays a crucial role in the dynamic adjustment

of context, with its influence on the familiarity signal being more indirect, occurring through its impact on C_t. Consequently, ε's influence on BOLD signal might be subtle and challenging to detect throughout the brain. This subtlety could perhaps be better captured using methods with higher temporal resolution, such as magnetoencephalography, or with improved spatial resolution, such as laminar fMRI (Stephan et al., 2019).

To summarize, our findings highlight that an important functional property of the AI is to update the level of familiarity of a given stimulus while also adapting to task-relevant, contextual information, processing expected information and expectations about what is likely to occur with heightened weight (Summerfield and de Lange, 2014). Ultimately, these expectations, combined with input visual signals through reciprocally interconnected feedback and feedforward processes, facilitate recognition judgments (Ploran et al., 2007; Kahnt et al., 2011; Summerfield and de Lange, 2014; Weilnhammer et al., 2018).

Lastly, while the laterality of insular function remains not fully understood (Craig, 2009; Lamichhane et al., 2016), bilateral insular activation has been reported in both recognition (Ploran et al., 2007) and perceptual decision-making (Thielscher and Pessoa, 2007), consistent with our findings.

It is worth noting that, according to the face processing extended system outlined by Haxby et al. (2000), the insula is linked to the processing of emotional expressions and emotional content associated with familiar facial stimuli (Haxby et al., 2000; Gobbini and Haxby, 2006, 2007; Zhang et al., 2018). This could potentially represent a possible alternative explanation of the observed AI activity. However, in the context of the face-recognition task utilized by Zaragoza-Jimenez et al. (2023), only neutral, computer-generated faces were employed. Therefore, it is unlikely that our findings reflect stimulus-related emotional response, as the insula shows an increased response for faces with whom one has a particularly intense emotional relationship (Gobbini and Haxby, 2006, 2007).

About the PI and face processing regions

The insula is a relatively large substrate whose subregions, such as the anterior and posterior portions, exhibit different functional properties (Zhao et al., 2023). The role of the PI is mainly associated with pain (Wiech et al., 2014) and sensorimotor processing (Craig, 2002; Kurth et al., 2010). Interoceptive signals initially reach this region bilaterally to undergo processing of low-level sensory features (Zhao et al., 2023). Then, this information is relayed to the AI, where higher-order interoceptive representations reach awareness (Craig, 2002; Zhao et al., 2023), and perceptual decisions are made. This process is supported, for example, by Wiech et al. (2010) who showed that, within the context of pain perception, the AI, in conjunction with the midcingulate cortex, integrates information regarding the significance of an impending stimulation into perceptual decision-making. Therefore, the PI serves as a gateway facilitating the transfer to higher-order processing without directly engaging in predictive coding processes. Considering its prominent role in pain and sensorimotor processing, we did not expect activity within the PI to reflect any of the predictive coding quantities described in the present study. However, the results showed BOLD percentage signal changes significantly lower than zero within this region, both in the left and right hemispheres. In particular, the left PI exhibited a significant below zero decrease in BOLD percentage signal change as a function of $VI_{i,t}$, while the right PI showed similar decreases in BOLD percentage signal change as a function of $VI_{i,t}$, δ, C_t, and F_t. These negative associations are noteworthy as they likely indicate processes of functional inhibition or downregulation (Wilson et al., 2020), suggesting that this insular subportion may indeed not be necessary for task performance in the present context. However, it is important to recognize that characterizing negative BOLD responses remains a challenging task. While these responses are thought to reflect cortical deactivation to some extent, they do not simply represent the inverse of positive BOLD responses (Huber et al., 2014; Mullinger et al., 2014; Wilson et al., 2020), as demonstrated, for instance, by Bressler et al. (2007) who showed that negative BOLD responses resulting from visual stimulation can carry high levels of stimulus-specific information.

In analyzing the core face system regions, our aim was not to reproduce Zaragoza-Jimenez et al. (2023) results, as the construction of design matrixes, exact ROIs localization and their definition (i.e., functionally defined in Zaragoza-Jimenez et al., 2023; anatomically defined in the present study), conducted statistical analyses, and participants included differed from those in the aforementioned study. Rather, the aim was to assess the robustness of the effects identified in Zaragoza-Jimenez et al. (2023). The methodological and analytical differences explain why, although the results demonstrate the same consistent overall involvement of the FFA and OFA, and not the pSTS, there are some minor discrepancies. Specifically, while our study reveals modulation of activity in both the right and left FFA as a function of $VI_{i,t}$, δ, C_t, and F_t, Zaragoza-Jimenez et al. (2023) observe this effect only in the right region. Regarding the OFA, where both studies identify a significant effect only in the right region, Zaragoza-Jimenez et al. (2023) observe an effect solely as a function of δ, whereas in the current study, it is observed as a function of $VI_{i,t}$, δ, C_t, and F_t. The core face system, described as a distributed hierarchical cortical network (Haxby et al., 2000; Duchaine and Yovel, 2015), is renowned to play a crucial role for the recognition of familiar faces and the encoding of invariant facial features (Pitcher et al., 2011a,b). This network is also involved in processing abstract high-level features of faces critical for identity recognition (Kanwisher and Yovel, 2006) and perceiving changeable facial features, like expression and eye gaze (O'Toole et al., 2002). Consistent with the findings of Zaragoza-Jimenez et al. (2023), our results support the role of the FFA and OFA, but not the pSTS, in contributing to these processes through the computing of predictive coding quantities that update feature-level familiarity and integrate contextual information, consistent with the view that familiarity with faces can be acquired regardless of the viewpoint from which they are perceived (i.e., view-independent processing). However, the decision of whether a face will be recognized also depends on the familiarity of recently perceived faces, emphasizing the role of contextual information in face identity recognition (Apps and Tsakiris, 2013; Zaragoza-Jimenez et al., 2023).

For a comprehensive discussion on the role of these areas in identity recognition and predictive coding, we direct the reader to the works of Zaragoza-Jimenez et al. (2023) and Apps and Tsakiris (2013).

Study limitations

This study is not devoid of potential limitations that necessitate consideration. As highlighted by Zaragoza-Jimenez et al. (2023) and discussed in Results, Face identity learning task of the present study, the computational parameters partly covary across trials, thus making it challenging to attribute single-parameter–based significance solely to the unique contribution of that parameter. Additionally, while our findings provide insights

into the role of predictive coding in inference and learning processes, it is important to acknowledge that examining brain area activations alone does not definitively corroborate the plausibility of predictive coding as a general theory (Stephan et al., 2019). To address this, laminar fMRI represents a promising method for directly testing such proposals in humans. Nevertheless, even with this approach, studying the insula poses challenges due to its complex folding and proximity to other brain structures.

References

Allen M, Fardo F, Dietz MJ, Hillebrandt H, Friston KJ, Rees G, Roepstorff A (2016) Anterior insula coordinates hierarchical processing of tactile mismatch responses. Neuroimage 127:34–43.

Amunts K, Mohlberg H, Bludau S, Zilles K (2020) Julich-Brain: a 3D probabilistic atlas of the human brain's cytoarchitecture. Science 369:988–992.

Apps MA, Tsakiris M (2013) Predictive codes of familiarity and context during the perceptual learning of facial identities. Nat Commun 4:2698.

Barrett LF, Simmons WK (2015) Interoceptive predictions in the brain. Nat Rev Neurosci 16:419–429.

Bastin J, et al. (2016) Direct recordings from human anterior insula reveal its leading role within the error-monitoring network. Cereb Cortex 27:1545–1557.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. Neuron 76:695–711.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.

Billeke P, Ossandon T, Perrone-Bertolotti M, Kahane P, Bastin J, Jerbi K, Lachaux JP, Fuentealba P (2020) Human anterior insula encodes performance feedback and relays prediction error to the medial prefrontal cortex. Cereb Cortex 30:4011–4025.

Bressler D, Spotswood N, Whitney D (2007) Negative BOLD fMRI response in the visual cortex carries precise stimulus-specific information. PLoS One 2:e410.

Castelhano J, Duarte IC, Wibral M, Rodriguez E, Castelo-Branco M (2014) The dual facet of gamma oscillations: separate visual and decision making circuits as revealed by simultaneous EEG/fMRI. Hum Brain Mapp 35:5219–5235.

Chand GB, Dhamala M (2016) The salience network dynamics in perceptual decision-making. Neuroimage 134:85–93.

Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci 36:181–204.

Corlett PR, Mollick JA, Kober H (2022) Meta-analysis of human prediction error for incentives, perception, cognition, and action. Neuropsychopharmacology 47:1339–1349.

Craig AD (2002) How do you feel? Interoception: the sense of the physiological condition of the body. Nat Rev Neurosci 3:655–666.

Craig AD (2009) How do you feel—now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70.

Crottaz-Herbette S, Menon V (2006) Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. J Cogn Neurosci 18:766–780.

Downar J, Crawley AP, Mikulis DJ, Davis KD (2001) The effect of task relevance on the cortical response to changes in visual and auditory stimuli: an event-related fMRI study. Neuroimage 14:1256–1267.

Duchaine B, Yovel G (2015) A revised neural framework for face processing. Annu Rev Vis Sci 1:393–416.

Dunovan KE, Tremel JJ, Wheeler ME (2014) Prior probability and feature predictability interactively bias perceptual decisions. Neuropsychologia 61:210–221.

Fazeli S, Büchel C (2018) Pain-related expectation and prediction error signals in the anterior insula are not related to aversiveness. J Neurosci 38:6461–6474.

Ficco L, Mancuso L, Manuello J, Teneggi A, Liloia D, Duca S, Costa T, Kovacs GZ, Cauda F (2021) Disentangling predictive processing in the brain: a meta-analytic study in favour of a predictive network. Sci Rep 11:16258.

FitzGerald TH, Moran RJ, Friston KJ, Dolan RJ (2015) Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. Neuroimage 107:219–228.

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360:815–836.

Friston K (2008) Hierarchical models in the brain. PLoS Comput Biol 4:e1000211.

Gładziejewski P (2016) Predictive coding and representationalism. Synthese 193:559–582.

Gobbini MI, Haxby JV (2006) Neural response to the visual familiarity of faces. Brain Res Bull 71:76–82.

Gobbini MI, Haxby JV (2007) Neural systems for recognition of familiar faces. Neuropsychologia 45:32–41.

Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63–72.

Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233.

Hesselmann G, Sadaghiani S, Friston KJ, Kleinschmidt A (2010) Predictive coding or evidence accumulation? False inference and neuronal fluctuations. PLoS One 5:e9926.

Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. Biom J 50:346–363.

Huber L, Goense J, Kennerley AJ, Ivanov D, Krieger SN, Lepsien J, Trampel R, Turner R, Möller HE (2014) Investigation of the neurovascular coupling in positive and negative BOLD responses in human brain at 7 T. Neuroimage 97:349–362.

Ishai A, Ungerleider LG, Haxby JV (2000) Distributed neural systems for the generation of visual images. Neuron 28:979–990.

Jia K, Xue X, Lee JH, Fang F, Zhang J, Li S (2018) Visual perceptual learning modulates decision network in the human brain: the evidence from psychophysics, modeling, and functional magnetic resonance imaging. J Vis 18:9.

Kahnt T, Grueschow M, Speck O, Haynes JD (2011) Perceptual learning and decision-making in human medial frontal cortex. Neuron 70:549–559.

Kanwisher N, Yovel G (2006) The fusiform face area: a cortical region specialized for the perception of faces. Philos Trans R Soc Lond B Biol Sci 361:2109–2128.

Keller GB, Mrsic-Flogel TD (2018) Predictive processing: a canonical cortical computation. Neuron 100:424–435.

Kok P, Failing MF, de Lange FP (2014) Prior expectations evoke stimulus templates in the primary visual cortex. J Cogn Neurosci 26:1546–1554.

Kok P, Mostert P, De Lange FP (2017) Prior expectations induce prestimulus sensory templates. Proc Natl Acad Sci U S A 114:10473–10478.

Kurth F, Zilles K, Fox PT, Laird AR, Eickhoff SB (2010) A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. Brain Struct Funct 214:519–534.

Lamichhane B, Adhikari BM, Dhamala M (2016) The activity in the anterior insulae is modulated by perceptual decision-making difficulty. Neuroscience 327:79–94.

Li S, Mayhew SD, Kourtzi Z (2009) Learning shapes the representation of behavioral choice in the human brain. Neuron 62:441–452.

Liu T, Pleskac TJ (2011) Neural correlates of evidence accumulation in a perceptual decision task. J Neurophysiol 106:2383–2398.

Loued-Khenissi L, Pfeuffer A, Einhäuser W, Preuschoff K (2020) Anterior insula reflects surprise in value-based decision-making and perception. Neuroimage 210:116549.

Luce RD (1977) The choice axiom after twenty years. J Math Psychol 15:215–233.

Luce RD, Ng CT, Marley AA J, Aczél J (2008) Utility of gambling I: entropy modified linear weighted utility. Econ Theory 36:1–33.

Meaux E, Sterpenich V, Vuilleumier P (2019) Emotional learning promotes perceptual predictions by remodeling stimulus representation in visual cortex. Sci Rep 9:16867.

Menon V, Uddin LQ (2010) Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214:655–667.

Mostert P, Kok P, de Lange FP (2015) Dissociating sensory from decision processes in human perceptual decision making. Sci Rep 5:18253.

Mullinger KJ, Mayhew SD, Bagshaw AP, Bowtell R, Francis ST (2014) Evidence that the negative BOLD response is neuronal in origin: a simultaneous EEG–BOLD–CBF study in humans. Neuroimage 94:263–274.

O'Toole AJ, Roark DA, Abdi H (2002) Recognizing moving faces: a psychological and neural synthesis. Trends Cogn Sci 6:261–266.

Perry CJ, Fallah M (2014) Feature integration and object representations along the dorsal stream visual hierarchy. Front Comput Neurosci 8:84.

Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N (2011a) Differential selectivity for dynamic versus static information in face-selective cortical regions. Neuroimage 56:2356–2363.

Pitcher D, Walsh V, Duchaine B (2011b) The role of the occipital face area in the cortical face perception network. Exp Brain Res 209:481–493.
Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME (2007) Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. J Neurosci 27:11912–11924.
R Core Team (2023) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/
Robinson D, Hayes A, Couch S (2023). Broom: convert statistical objects into tidy tibbles. R package version 1.0.5. Available at: https://CRAN.R-project.org/package=broom
Sayal A, Sousa T, Duarte JV, Costa GN, Martins R, Castelo-Branco M (2020) Identification of competing neural mechanisms underlying positive and negative perceptual hysteresis in the human visual system. Neuroimage 221:117153.
Seeley WW (2019) The salience network: a neural system for perceiving and responding to homeostatic demands. J Neurosci 39:9878–9882.
Seth AK (2013) Interoceptive inference, emotion, and the embodied self. Trends Cogn Sci 17:565–573.
Siman-Tov T, Granot RY, Shany O, Singer N, Hendler T, Gordon CR (2019) Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. Neurosci Biobehav Rev 105:262–275.
Stephan KE, Petzschner FH, Kasper L, Bayer J, Wellstein KV, Stefanics G, Pruessmann KP, Heinzle J (2019) Laminar fMRI and computational theories of brain function. Neuroimage 197:699–706.
Sterzer P, Kleinschmidt A (2010) Anterior insula activations in perceptual paradigms: often observed but barely understood. Brain Struct Funct 214:611–622.
Sterzer P, Voss M, Schlagenhauf F, Heinz A (2019) Decision-making in schizophrenia: a predictive-coding perspective. Neuroimage 190:133–143.
Summerfield C, De Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. Nat Rev Neurosci 15:745–756.
Thielscher A, Pessoa L (2007) Neural correlates of perceptual choice and decision making during fear–disgust discrimination. J Neurosci 27:2908–2917.
Todorov A, Said CP, Engell AD, Oosterhof NN (2008) Understanding evaluation of faces on social dimensions. Trends Cogn Sci 12:455–460.
Verdade A, Sousa T, Castelhano J, Castelo-Branco M (2022) Positive hysteresis in emotion recognition: face processing visual regions are involved in perceptual persistence, which mediates interactions between anterior insula and medial prefrontal cortex. Cogn Affect Behav Neurosci 22:1275–1289.
Walsh KS, McGovern DP, Clark A, O'Connell RG (2020) Evaluating the neurophysiological evidence for predictive processing as a model of perception. Ann N Y Acad Sci 1464:242–268.
Watanabe T, Sasaki Y (2015) Perceptual learning: toward a comprehensive theory. Annu Rev Psychol 66:197–221.
Weilnhammer VA, Stuke H, Sterzer P, Schmack K (2018) The neural correlates of hierarchical predictions for perceptual decisions. J Neurosci 38:5008–5021.
Wickham H, François R, Henry L, Müller K, Vaughan D (2023). dplyr: A grammar of data manipulation. R package version 1.1.2. Available at: https://CRAN.R-project.org/package=dplyr
Wickham H, Vaughan D, Girlich M (2024). tidyr: Tidy Messy data_. R package version 1.3.1. Available at: https://CRAN.R-project.org/package=tidyr
Wiech K, Jbabdi S, Lin CS, Andersson J, Tracey I (2014) Differential structural and resting state connectivity between insular subdivisions and other pain-related brain regions. Pain 155:2047–2055.
Wiech K, Lin CS, Brodersen KH, Bingel U, Ploner M, Tracey I (2010) Anterior insula integrates information about salience into perceptual decisions about pain. J Neurosci 30:16324–16331.
Wilson R, Thomas A, Mayhew SD (2020) Spatially congruent negative BOLD responses to different stimuli do not summate in visual cortex. Neuroimage 218:116891.
Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM (2004) Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage 21:1732–1747.
Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14:1370–1386.
Zaragoza-Jimenez N, Niehaus H, Thome I, Vogelbacher C, Ende G, Kamp-Becker I, Endres D, Jansen A (2023) Modeling face recognition in the predictive coding framework: a combined computational modeling and functional imaging study. Cortex 168:203–225.
Zhang Y, Zhou W, Wang S, Zhou Q, Wang H, Zhang B, Huang J, Hong B, Wang X (2018) The roles of subdivisions of human insula in emotion perception and auditory processing. Cereb Cortex 29:1–12.
Zhao H, Turel O, Bechara A, He Q (2023) How distinct functional insular subdivisions mediate interacting neurocognitive systems. Cereb Cortex 33:1739–1751.