Disentangling Predictive Processing in the Brain: A Meta-analytic Study in Favour of a Predictive Network

Linda Ficco (linda.ficco@uni-jena.de)
Friedrich-Schiller University of Jena

Lorenzo Mancuso
University of Turin

Jordi Manuello
University of Turin

Alessia Teneggi
University of Turin

Donato Liloia
University of Turin

Sergio Duca
GCS-fMRI, Koelliker Hospital

Tommaso Costa
University of Turin

Gyula Zoltán Kovacs
Friedrich-Schiller University of Jena

Franco Cauda
University of Turin

Research Article

Keywords: predictive coding (PC) theory, brain, ctivation Likelihood Estimation (ALE) algorithm

Posted Date: March 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-296410/v1

Version of Record: A version of this preprint was published at Scientific Reports on August 10th, 2021.
See the published version at https://doi.org/10.1038/s41598-021-95603-5.

Abstract

According to the predictive coding (PC) theory, the brain is constantly engaged in predicting its upcoming states and refining these predictions through error signal. Despite extensive research has investigated the neural bases of this theory, to date no previous study has systematically attempted to define the neural mechanisms of predictive coding across studies and sensory channels, focussing on functional connectivity. In this study, we employ a coordinate-based meta-analytical approach to address this issue. We first use the Activation Likelihood Estimation (ALE) algorithm to detect spatial convergence across studies, related to prediction error and encoding. Overall, our ALE results suggest the ultimate role of the left inferior frontal gyrus and left insula in both processes. Moreover, we employ a task-based meta-analytic connectivity method (Seed-Voxel Correlations Consensus). This technique reveals a large, bilateral *predictive network*, which resembles large-scale networks involved in task-driven attention and execution. In sum, we find that: i) predictive processing seems to occur more in certain brain regions than others, when considering different sensory modalities at a time; ii) there is no evidence, at the network level, for a distinction between error and prediction processing.

1. Introduction

Both when we are resting and when we are involved in tasks, our brain attempts to “figure out”, at many levels, what the immediate as well as the subsequent future events are going to look like. This is the main assumption of the predictive coding (PC) theory [1–5]. PC theory received extensive support recently from a vast range of theoretical and experimental studies, regarding primary sensory processes [6], language [7] and higher level cognitive processes, such as decision making and naturalistic speech comprehension [8,9]. Moreover, evidences have been obtained with a variety of methods, mostly by functional magnetic resonance imaging (fMRI), but also by electroencephalography [10,11], computational simulations [12], physiological recordings of single neurons (for a review, see [13], as well as transcranial magnetic stimulation [14].

According to this theory, the brain is a hierarchically organized system where, at each level of processing, higher layers try to “predict” the latent causes of the sensory inputs coming from lower layers. In fact, top down-modulation of prior expectations has been documented in both low-level [5,15] and high-level processing [16,17]. A key element of the theory is the possibility for the brain to learn, i.e. to correctly update its internal models of the world - through prediction error signals [18–20]. Prediction errors emerge when there is a discrepancy between what is expected at a certain level and the actual sensory information coming from the lower processing stages. Since 1999, when Rao and Ballard published their seminal simulation work, the number of attempts to implement PC in the human brain have increased sharply. It has been argued that predictive processing occurs at a cellular level [21], where the activity of neural populations is modulated by higher-order predictions and units signalling precision of these predictions. According to [22]Bastos and colleagues, PC is a typical property of the human cerebral neocortex because its layered structure would suit a hierarchical signal exchange between cortical layers. In particular, error

signals seem to be computed in the granular layers (especially layer IV), while predictions would be encoded in layers II and III. Friston and colleagues hypothesized a large set of brain areas with predictive processing mechanisms, including the primary sensory and motor cortices, motor association cortices, dorsal and ventral prefrontal cortices, parietal cortex, anterior cingulate cortex, insula, hippocampus, amygdala, basal ganglia, thalamus, hypothalamus, cerebellum and the superior colliculus [23,24]. In a recent review Keller and Mrsic-Flogel pointed out the necessity of the functional separation between neuronal units computing predictions and the ones involved in the generation of the error signals [25]. This separation has been found empirically in a mathematical modelling of the auditory cortex, where neurons encoding predictions were located in cortical layers II/III and prediction error neurons in layer IV [12]. The hypothesized distinction led to the idea of separating cortical functions related to a violation of predictions from those that create, maintain and update them.

Moreover, the identification of a brain network, specialized in the encoding of predictions and transmission of error signal across sensory modalities at a large-scale level is a rather open and unexplored field. In fact, while earlier formulations of PC theory ground such mechanisms in different layers of the human cortex, more recent models of PC attribute functions of error computation and prediction encoding to discrete brain regions and their interconnections [26–28]. In line with a network-like view of brain function [29], the present research work aims at investigating the potential presence of a specific, predictive network. To our knowledge, no previous study has approached the definition of such a network, adopting meta-analytic functional connectivity methods. To do so, we employed a coordinate-based meta-analytic approach, utilising the Activation Likelihood Estimation (ALE) technique, and adopting functional connectivity methods previously used in clinical neuroimaging [30]. We expected the following results:

a) In general, at least some of the regions found in the literature to be involved in predictive activity might be found to be functionally connected, revealing a spatially defined network.

b) For the brain areas generally involved in predictive processing, we might only partially replicate the results from a recent meta-analysis [31], principally due to the diversity of data collection and organization criteria. Moreover, given the dense interchange of prediction and error signals in the human cortex [22,25], and the heterogeneous nature of the experimental tasks included in our datasets, we also hypothesized that mostly higher order regions shall be found as the common neural correlates of PC.

c) A recent meta-analysis [18] highlighted the contribution of striatum, insula, thalamus and fronto-medial structures to prediction error signal in a recent meta-analysis, while others [32] reported other regions (the bilateral ventral striatum, the thalamus, the left frontal operculum, the left caudate and the left IFG). We

therefore aimed at verifying if, with different selection and categorization criteria, these results on prediction error computation could be replicated.

2. Results

2.1 Selection of studies

By following the criteria (a-f) described in Sect. 5.1, 106 articles were collected (see Fig. 1). Data from these articles were classified in a table specifying an identification code, year of publication, first author's family name, title, scientific journal, number of experimental subjects, experimental task, sensory modality investigated, experimental contrast, type of stimuli. A further selection based on criteria g) and h) led to 70 articles. Those works which were discarded at this point were included in a separate list, along with the reason for their exclusion. All the peak coordinates listed for the experimental contrast, which were classified as Prediction Encoding or Prediction Violation, were listed in a separate table. The classification of each reported contrast can be found in the Supplementary Tables S1 and S2. When necessary, we converted the peak coordinates to the Montreal Neurological Institute (MNI) space, using the icbm_spm2tal transform on GingerALE [37,42], see http://www.brainmap.org/icbm2tal/.

2.2 Activation Likelihood Estimation

As a first step, we conducted a general ALE analysis across contrasts on the results from all the selected articles (70 experiments, 930 foci of activation and 1419 participants). We refer to these analyses as General Prediction. Afterwards, the general dataset was further divided into two conditions: Prediction Violation (45 experiments, 511 foci and 939 subjects) and Prediction Encoding (39 experiments, 444 foci, 750 subjects). We performed ALE meta-analyses singularly on these two conditions as well. Figure 2 shows the results of the ALE analyses at FWE, $p < 0.05$. Further details of the ALE results are reported in Table 1.

Table 1
Activation likelihood estimation (ALE) results. Convergent findings of brain activity related to predictive coding conditions.

Condition	MNI coordinates (x,y,z)	Volume (mm³)	Maximum ALE value	Z score	P value	Anatomical location (Brodmann area)
General Prediction*						
	-46, 10, 24	1088	0.049	6.78	6.11e-12	Left inferior frontal gyrus (BA 9)
	-30, 24, 2	608	0.043	6.10	5.23e-10	Left insula (BA 13)
	4, 14, 50	224	0.037	5.39	3.59e-8	Right superior frontal gyrus (BA 6)
	34, 24, 2	200	0.034	5.09	1.83e-7	Right insula (BA 13)
	-28, -64, 46	200	0.034	5.13	1.44e-7	Left precuneus (BA 19)
Prediction Encoding**						
	36, -60, 50	832	0.021	4.47	3.96e-6	Right superior parietal lobule (BA 7)
	-50, -44, 52	608	0.021	4.38	5.82e-6	Left inferior parietal lobule (BA 40)
	4, 12, 50	584	0.026	4.98	3.19e-7	Right Superior frontal gyrus (BA 6)
	-38, -80, -14	528	0.023	4.67	1.49e-6	Left fusiform gyrus (BA 19)
	48, -64, -8	464	0.018	3.98	3.44e-5	Right fusiform gyrus (BA 19)
	-46, 8, 26	440	0.023	4.61	2.03e-6	Left inferior frontal gyrus (BA 9)
	48, -44, 46	328	0.020	4.19	1.41e-5	Right inferior parietal lobule (BA 40)
	28, -6, -20	312	0.021	4.35	6.70e-6	Right amygdala
	42, 22, 32	288	0.019	4.14	1.76e-5	Right middle frontal gyrus (BA 9)
	-28, -6, -20	224	0.020	4.25	1.05e-5	Left amygdala
	-26, -64, 46	224	0.019	3.98	3.40e-5	Left precuneus (BA 7)
	46, 18, 6	192	0.021	4.29	8.81e-6	Right insula (BA 13)
	20, -100, 6	136	0.018	3.82	6.80e-5	Right cuneus (BA 17)

Condition	MNI coordinates (x,y,z)	Volume (mm³)	Maximum ALE value	Z score	P value	Anatomical location (Brodmann area)
	66, -22, 6	120	0.017	3.79	7.41e-5	Right superior temporal gyrus (BA 42)
Prediction Violation*						
	-46, 10, 24	616	0.033	5.59	1.12e-8	Left inferior frontal gyrus (BA 9)
	-30, 24, 0	608	0.038	6.19	2.94Ee-10	Left insula (BA 13)

***** Significant activations are set at voxel-level p < 0.05 with the Family-wise error (FWE) correction. ****** Significant activations are set at p < 0.0005 uncorrected for multiple comparisons.

2.2.1 Prediction Violation

Two significant clusters were related to the violation of predictions (Fig. 2, red color). The larger cluster included the left inferior frontal gyrus, while a smaller one was found over the left anterior insular cortex, partially overlapping with the claustrum.

2.2.2 Prediction Encoding

No cluster was significant for Prediction Encoding at the typically applied, conservative threshold of FWE p < 0.05. Lowering the threshold to FDR p < 0.01 still did not produce any significant clusters. However, at an exploratory level, we report the results obtained at a more liberal threshold (*Uncorrected, p < 0.0005*). At this threshold, fourteen clusters emerged. These included the right superior and left inferior parietal lobules, the right superior, right middle and left inferior frontal gyri, the bilateral fusiform gyri and the right amygdala, and a few clusters with a size inferior to 200 mm^3 (including the left amygdala, the left precuneus and the right cuneus, the right insula and the right superior temporal gyrus).

2.2.3 General Prediction

Overall, the ALE analysis of the whole dataset returned a set of cortical regions in the frontal and parietal lobes (Fig. 2, green). These include the left inferior frontal gyrus, the insulae bilaterally, the right superior frontal gyrus, the bilateral inferior parietal lobules, and the left precuneus.

2.3 Seed-Voxel Correlations Consensus

This technique highlights the regions showing correlated activity with those that were active during the tasks tapping into predictive processing (see Sect. 5.3). Overall, the results from all the three conditions (Prediction Encoding, Violation and General Prediction) are remarkably similar, and they involve a bilateral set of brain regions. Significant peaks are located in the left inferior frontal gyrus, the superior temporal gyrus bilaterally, the left thalamus, the left hippocampus and the left cerebellum, and significant voxels

are shown in warm colors. The network emerging from negative correlations is shown in cold colors, and it includes the right cerebellum (uvula), the left precentral gyrus and the post-central gyri bilaterally, and the right middle occipital gyrus (Fig. 3; see Table S2 for more details). Finally, the network relative to Encoding is substantially overlapping with that of the other two conditions, although the map of positive values appeared to be less extended (especially in the middle cingulate gyrus, the insula and the cerebellum). This may indicate relatively lower interconnections to areas pertaining error detection and salience.

Although the SVC Consensus maps indicate regions that are significantly connected to the activation foci reported in the literature, the relation between these foci and such maps needs to be clarified. In fact, on one hand it is possible that only a few foci are responsible for the connectivity maps. On the other hand, these maps might show areas that are not reported by the literature (thus are not primarily considered to be involved in predictive processing), but are systematically connected to the predictive regions, possibly providing input or output to them. To investigate the relation between the foci and their connectivity, we overlapped the SVC consensus of the General Prediction condition to the corresponding unthresholded ALE map. Here, the unthresholded map can be seen as an indicator of all the activated regions in the literature. We found out that there is a substantial overlap between the two maps (Fig. 4), suggesting that the activated areas tend to be interconnected and to form a coherent functional network (Cosine Similarity Index = 0,56). Lastly, to exclude the possibility of bias due to local connectivity, the SVC Consensus analysis was repeated excluding proximal connections between close areas from the SVC maps (see Sect. 5.3). The resulting maps were extremely similar to the original SVC consensus maps, suggesting that activation foci are connected not only to the spatially closer areas, but also to the more distal ones (Figure S1).

2.4 Fail-safe technique

To test the impact of potentially missing data on our results, the fail-safe method [40] has been applied. In the General Prediction condition, the analysis shows that at least one of the clusters remains significant up to the infusion of 250% random data (Fig. 5). The analysis of the Prediction Violation dataset suggests the stability of the data as well (the clusters remain significant up to the inclusion of 425% random data). In general, both fail-safe tests suggest the robustness of the two clusters that are in common for the two conditions, i.e. the left IFG/precentral gyrus and the left insula/claustrum.

2.5 Leave-N-out

This analysis tests whether all the studies in a dataset contribute to the results similarly, or, in other words, whether a dataset is homogeneous. Figure 6 shows the results from the Leave-n-out analysis. The y axis indicates the number of papers, while the x axis the energy (1-quadratic error/total N experiments, see Methods Sect. 5.4). The diagrams show the distribution of energy obtained by removing 3,5,7,9 and 11 articles at each run separately. When removing less than 7 random articles at a time, no important changes are visible in the distribution. Since 7/39 (removed articles/total) equals to 18% of the included

experiments, this suggests that the condition of Prediction Encoding is mostly homogeneous. Thus, the absence of significant convergence may be due to the heterogeneity of the activation coordinates, and is unlikely to be due to the heterogeneity in the experiments *per se*.

3. Discussion

The ALE results show convergence across tasks targeting predictive processing in a set of cortical regions, both in the Violation and in the General condition. However, we were also unable to detect convergence in the Encoding condition. This suggests that the encoding of predictions happens across the brain in a spatially distributed fashion, without involving specific areas. As suggested by the results of the Leave-N-Out procedure, this spatial heterogeneity seems not to be due to the disproportionate contribution of few outlier studies, rather to the large variability of the localization of foci that characterize the entire condition. The relatively homogeneous spatial distribution of prediction encoding activity is not surprising. In fact, there would be no theoretical need for discrete brain regions dedicated to this step, because expectations about the hidden causes of a sensory input are coded in units located evenly in specific cortical layers [43]. Further, this result fits with the idea, pivotal in PC theorizations [17], that perhaps the most important function of a perceptual system is producing an error signal when a mismatch between current representations and sensory input occurs. In this sense, the only signal that our convergence analysis could detect would thereby be the concerted activity of prediction error units [44]. Additionally, it should be noted that the current spatial resolution capacity of the typical fMRI scanner does not allow the investigation of the different activity patterns across cortical layers, which would indeed be more revealing in this context.

Instead, an expected location of convergence that we did not detect is the cerebellum, especially when analysing the violations of predictions. In fact, this brain structure has been reported in previous works as an important hub processing the comparison between an internal model and the current sensory input, and as a region supporting procedural and perceptual learning mechanisms [45–47]. However, according to a recent study [48], there are two main reasons why not many neuroimaging meta-analyses are able to detect convergence in the cerebellum. First, there would be technical difficulties related to MRI signal detection from the cerebellum, principally dependent on experiments targeting climbing fibres, which are poorly coupled to the BOLD signal [48]. Second, some experimental paradigms tend to promote rapid habituation within the area, again eliciting less BOLD signal [48].

Our analyses of the Prediction Violation condition confirm previous findings (see [18,32]) as regards the insula and the IFG, while the involvement of others, such as the striatum and the thalamus, are not confirmed. Both the insula and the IFG have been related to the violation of predictions in previous works. In particular, the role of insular regions in prediction error computation is well-known already, especially in the interoceptive modality [28,49,50]. Thus, it seems that, when considering error signals cross-modally, these tend to converge in regions that process self-related information. Considering that only a small number of included studies targeted bodily sensations explicitly, this finding deserves special attention. A

tentative interpretation is that, regardless of the nature of the specific expectations that are violated in each task, these tend to produce an error broadly related to the self. The insular cluster also extends to include the left claustrum, which produces prediction error signals in Pavlovian classical conditioning paradigms [51]. Notably, these involve a component of automatic learning, probably present in the majority of the tasks included here to a certain extent. The second cluster was located in the IFG, which is involved in risk aversion [52], and in detecting a mismatch between expectations and decisions [53]. Interestingly, results from a computational study found a correlation between both IFG and insula activity with a prediction error model during bi-stable perception, which is a paradigm inducing strong violations of visual expectations [54]. Moreover, intrinsic connectivity between the IFG and the insula was predicted by the degree of intolerance of uncertainty, which indicates their sensitivity to error signals [55]. Together, these findings and our results suggest that the insula and the IFG, and their connectivity during prediction violation across modalities is a worthy avenue for further research.

The General Prediction condition was designed to tap into the general effects of predictive processing, resulting from the mere fact of performing a task eliciting predictions or prediction errors, more or less directly. Thus, we expected the brain regions emerging here to be related either to one or the other process, or to both. First, convergence was found in two larger clusters, one in the inferior frontal gyrus/precentral gyrus, and the other bilaterally in the insula. Strikingly, both regions also emerged in the analysis of the general predictive condition in a similarly constructed recent meta-analysis [31], strengthening the plausibility of this result. In further support of the double role of both the IFG and the insula in PC, activity in these areas represents the building of an expectation, analysing the conjunction across somatosensory, visual and auditory stimulus modality [56]. However, evidence about the IFG in particular is somewhat mixed. In fact, not always its activity seems to depend on the predictability of a situation [57]. Moreover, whereas some authors describe it as an area involved in the processing of “expectancy input” [58]others report increased activity in the IFG activity when the stimulus probability is low, leading to larger prediction error signal [59]. As regards the anterior insulae, notably these are an important hub of the salience network [60]. This hub is plausibly more activated in surprising situations driving attention, which also involve an increased gain in error signal computation. The relationship between predictive and attentional processes is extensively discussed in the following Sect. (4.3). Lastly, the role of precuneus in the General Prediction condition is more difficult to relate to existing literature. In general, this structure is involved in self-related cognition, episodic memory and mental imagery [61]. Interestingly, this region was responsive to deviant stimuli even during sleep [62], which may indicate a selective sensitivity to prediction error during different states of consciousness. Overall, these studies support our findings, and suggest that the IFG and the insula might be involved both in the encoding and in the violation of predictions. More evidence on prediction violation than encoding exists in both cases, and the sensitivity of the IFG (and the less discussed precuneus) to stimulus probability may more strongly depend on specific task characteristics.

The SVCs Consensus analysis was conducted to highlight a set of interconnected regions that tends to co-activate with those involved in prediction violation and encoding. The resulting network is largely

similar between conditions, thus here we will only focus on the General Prediction condition. First of all, the wide overlap between this network and the unthresholded ALE maps indicates that the activated regions, despite being spatially heterogeneous (as no ALE convergence was found), nevertheless tend to be strongly interconnected during task execution.

Surprisingly, the maps relative to the Prediction Violation and Prediction Encoding conditions are almost completely overlapping. This seems to contradict evidence that prediction violation and encoding are functionally separated, as seen at a cellular [25] and at an infra-laminar level, between cortical layers [22,63]. Our ALE results further support this separation, since the two conditions generated different results. However, such separation may not exist at the network level. Indeed, the similar spatial extent of the two maps could simply mean that, during task execution, the regions that are involved in prediction cross-modally tend to exchange information with the same, broad set of areas. Actually, the map related to Prediction Encoding shows relatively lower spatial extension in some regions (Figure S2, insula, cerebellum, and anterior cingulate cortex), all involved in error computation [64,65]. Again, this difference in the overall map extension is in line with the hypothesis of stronger brain activity for error signal computation, rather than that of prediction encoding.

A core feature of this network is its remarkable similarity to the so-called task-positive network (TPN, [66])The TPN is a set of areas involved in task execution, and is usually divided into three large-scale brain networks related to salience processing [60] and the dorsal and the ventral attentional networks [67,68].

The fact that the regions which are more involved in prediction are also part of attentional networks is of key importance. In fact, an ever-increasing amount of evidences suggest that both attention and prediction support perception. On one hand, attention enhances the processing of relevant information and regulates the general cortical responsiveness [69–71]. On the other hand, prediction allows the brain to take prior information into account [72] and “anchors” attentional processing, meaning that computing predictions is necessary to subsequent attention orientation [73]. The overlap between our map and the TPN could be interpreted in several ways. First, despite being distinct processes [74–77], they may share a common neural territory. Since attention is a mechanism aiding error minimisation, by adjusting the “volume” of prediction error units [78]. Specifically, attention adjusts the precision of predictions via synaptic gain enhancement [79,80], leading to increased error signals. Considering that this network relies on the original coordinates of brain activations during task performance, and that the BOLD signal reported therein likely measures prediction errors, it is well possible that both attention to the actual stimuli and the prediction of future stimuli were just at work simultaneously. Second, although PC has been extensively reported in sensory areas, the activity of prediction and error computation might specifically involve attentional networks more than other brain regions, when multiple modalities are considered together. This has never been observed before because, for obvious practical reasons, only a limited set of modalities are investigated at a time, often in a rather constrained experimental environment. Finally, a third possibility is that the TPN emerged from our analyses merely for the effect of the engagement of participants in any attention-demanding task, and the selected contrasts do not reflect

predictive processing at all. It is difficult to rule out this possibility completely, as we did not analyse an arguably non-predictive control neutral condition (for the role of neutral conditions, see: https://psyarxiv.com/y7639/). Still, our findings are the results of a careful selection of appropriate neuroimaging contrasts from a large number of original studies, so the discussed overlap with the canonical attentional networks might be taken as an evidence for some form of relationship between the two processes. Moreover, this aligns with an increasing corpus of research considering prediction and attention as dissociable, but strongly interdependent processes (for empirical evidences, see [73,81,82]; for further readings see [83–86]).

Another interesting consideration is that, as our predictive network overlapped with the TPN, it appeared to be negatively correlated with the default mode network (DMN; for example, see Fig. 3, negative values). In fact, the DMN and the TPN are also considered to be "anticorrelated" [87], and possibly involved in different forms of cognition. In particular, the DMN typically results as more activated during rest and mind-wandering [88]. Considering recent work suggesting that the DMN creates and updates internal predictive models about the self [89], and that it is engaged when stimuli are temporally predictable [90], the striking absence of PC involvement in the DMN is rather unexpected, especially in the Encoding condition. Moreover, since the DMN is located at the extreme in a continuum of integration and hetero-modal functioning within the human connectome [91,92], it is even more surprising that it did not result from our functionally heterogeneous PC-related dataset. A possible reason for its absence could be that, under the hypothesis that the DMN is responsible for the integration of predictions in former internal models – acting in a sort of "autopilot mode" [93,94], almost no experimental paradigm, included in the current analysis, tested this kind of automatic activity. Indeed, given the functional nature of the DMN, the specific impact of prediction on this network is difficult to test experimentally. Techniques with high temporal resolution, computational and meta-analytic approaches to functional neuroimaging data can be valuable tools to investigate the role of the DMN in predictive processing in the future.

One first, potential limitation of the current work is that the selection, classification and coding of the articles was conducted manually by one author only. In order to reduce the risk of errors occurring in these phases, the coded dataset was, however, cross-checked by another author, independently. Moreover, a section of notes was included in the database with the aim to make the interpretation and selection processes more transparent, as suggested by recent guidelines (Müller et al. 2016). Another potential flaw in our database is caused by the heterogeneity of the definition of "prediction" across studies. In fact, the concepts of prediction, anticipation and expectation are often used interchangeably [95] and how they are operationalized in each study can potentially lead to confusion with other processes [72,96]. Lastly, given the strong presence of studies employing visual or audio-visual tasks could have also limit the validity of the current results (see Table S4). However, the absence of early visual areas in both the ALE and the SVC Consensus results may suggest that the impact of this imbalance is nonetheless limited.

4. Conclusions

In sum, we brought evidence in favour of the existence of a predictive network, to our knowledge for the first time. First, the ALE convergence analyses suggest that fMRI seems to be more suitable to the detection of error signals in the investigation on predictive processing. Taken together, these results point to the presence of a brain network as a possible candidate for the processing of predictions and error signals across sensory modalities. Against the hypothesis that predictive coding occurs typically in cortical tissues in the whole brain, it rather seems to be supported by a specific, spatially extended bilateral network, mainly related to attentional processes. Finally, although the separation between error, weighting and encoding unit is supported by or ALE results and previous works, it may vanish when large-scale functional connectivity is examined.

5. Methods

5.1 Selection of studies

We searched for publications (Pubmed, https://www.ncbi.nlm.nih.gov/pubmed/) up to January 2019. Articles were chosen using the keywords “predictive coding” AND “fMRI”, OR “functional magnetic imaging” OR “functional brain imaging” in the title or in the abstract of the articles. The decision to choose the only term “predictive coding” instead of a variety of related terms had two purposes: on one hand, to select only articles explicitly explaining their results under this framework; on the other hand, we did not include terms as “prediction error”, “expectation” or “Bayesian brain” so that articles describing the role of expectancy in psychology without referring to PC could be excluded.

In the initial selection stage, the following primary inclusion criteria were applied. We included studies:

a) which employed fMRI;

b) which provided the peak coordinates of significant activation in stereotactic brain space (MNI, TAL);

c) which were original experimental works. We excluded reviews or other meta-analytic studies;

d) which reported whole-brain analysis for the contrasts of interest (i.e., articles which were only based on a priori Regions of Interests –ROIs analyses were excluded);

e) which were based on a healthy adult sample;

f) which were written in English.

Additional, more specific criteria for inclusion were the following:

g) the articles had to explicitly support the PC theory, regardless of the sensory modality and the process investigated (i.e., only studies bringing evidence in favour of the framework were considered);

h) the articles had to include experimental contrasts which reflect the violation of a prediction or the creation, updating or maintenance of a predictive internal model. The first is often related to the

generation of a prediction error signal, for example comparing a deviant event or condition with a regular one, or fitting a statistical model designed to represent the same effect. The second case, in our study operationalized in the category of Prediction Encoding, typically included fMRI contrasts between a learned and an unfamiliar condition/event, or between an expected and an unexpected condition. Figure 1 shows a flowchart representing the described steps of the selection process. For a detailed list of the contrasts included in the study and the classification of their reflected effect, see Table S1 and S2;

To classify the contrasts included in each experiment, we inspected both the nature of the task and the interpretation of the effects that the authors themselves reported in the articles in the following way. If the coordinates pertained to a condition where participants saw a deviant stimulus, leading to potential surprise reaction in a stream of predicted (standard) stimuli and reflecting prediction errors [18], the selected coordinates were classified as locations of “prediction violation”. If the coordinates represented a condition where a repetitive stimulation was applied, leading to statistically-based predictions [33], they were labelled as locations of “prediction encoding”. Since it is frequent that coordinates related to the error or the encoding effect are reported in the same experiments from different contrasts, not the experiments themselves but the single activation coordinates of the given contrast were classified according to these two categories. For a complete list of the different paradigms or tasks included in each category see Table S1. Furthermore, this classification not only took the performed task into account, but it considered also the interpretation of the results that was provided by the authors. For instance, interpreting a reduced response in some brain areas, related to the repetition of stimuli is often interpreted as repetition suppression (RS), hence this effect would be included in our Prediction Encoding category (for a review on how PC may explain RS see [26]). Finally, although recently a point has been made to distinguish the effects of repetition and expectation suppression [34], note that we consider both of these effects under the category of Prediction Encoding.

5.2 Activation Likelihood Estimation

Activation Likelihood Estimation (ALE) is a meta-analytic technique, which detects areas of convergence across peak coordinates of significant activations from functional neuroimaging studies [35,36]. In short, the current version of this algorithm models a Gaussian kernel for each activation peak, considering these as fixed-effects within each study. The width of the kernels accounts for between-subject and between-lab variations, and it models on the number of participants in each study [35]. Then, one spatial uncertainty map is calculated for each study, unifying all the modelled peaks, where studies are treated as random effects. Afterwards, in order to test for statistical significance, the algorithm calculates an iterative comparison between the modelled maps and a null distribution, reflecting random spatial association of peaks between studies. Specifically, this distribution reflects the null hypothesis that the activation foci are randomly distributed in the brain, leading to convergence only by pure chance. Lastly, a correction for multiple comparisons is applied. We used GingerALE version 3.0 (http://www.brainmap.org/) to perform the above meta-analysis. The threshold for detecting significant activations was set at voxel-level $p < 0.05$, with 1000 permutations with the family-wise error correction method (FWE), and the analyses were

performed in the MNI152 coordinate space. Coordinates reported in TAL in the original study were converted using the icbm2tal transform prior to the ALE analysis [37].

5.3 Seed-voxel correlations Consensus

To investigate the connectivity patterns of the areas involved in predictive processing, we performed a Seed-Voxel Correlations (SVC) Consensus technique, adapted to functional data. This technique was originally developed by Boes and colleagues [30] to map the connectivity of brain lesions and it consists of overlapping several SVC maps, to verify if they tend to connect to a set of shared areas. The explicit aim of this method was to test if the spatial heterogeneity of brain lesion of a given deficit could be reduced to a common functional network [38]. Similarly, in the present work we aimed to evaluate the spatial variability of regions associated with predictive processing. Specifically, we hypothesized that the diverse activation foci which were reported in the literature might belong to a single brain network, and thus that they tend to be connected to each other. To do so, each peak that entered the ALE meta-analysis was searched in the Neurosynth resting-state database to obtain a functional connectivity map. Neurosynth (http://www.neurosynth.org) is an online database of functional meta-data which allows to easily obtain SVC maps calculated on the 1000 subjects of the Brain Genomics Superstruct Project (https://dataverse.harvard.edu/dataverse/GSP). Each SVC map was then considered as an individual subject in a second-level analysis. Then, the overlap of those maps was assessed by the means of a one-sample T-test on SPM12, (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) with a FWE- corrected threshold of $P < 0.05$. Both the positive and the negative contrasts were calculated, thus obtaining a map related to the shared positive correlations and one related to the negative ones. This analysis was carried out separately for the datasets of the three conditions (Violation, Encoding and General; for details see below).

It should be noted that functional connectivity is partly influenced by physical closeness, so that spatially closer voxels tend to be connected more strongly [39]. As we worked with a large number of foci, it may be argued that many of the seeds were close to each other, and thus their maps would display a high degree of overlap. Therefore, the SVC Consensus results could be biased because of the mere spatial closeness of a vast proportion of peaks. In other words, it could be that there was a significant overlap not because the regions are functionally connected, but simply because most of the seeds of connectivity maps were close to each other. To test this hypothesis, we repeated the SVC Consensus analysis of the General Prediction condition using only the distant connectivity of each seed, defined as all the voxels in each SVC map that were at least 14 mm^3 far from the seed [39]. Hence, local connectivity of a seed is the volume within the radius of 14 mm around it. For this test, all the voxels around the seed in each of the SVC maps were set to 0, before recalculating the t-test.

5.4 Fail-safe technique

The fail-safe technique [40] allows the verification of the extent to which neuroimaging results resist the addition of noisy data and, generally, it is an indicator of potential selection bias in the dataset. To perform this analysis, different ALE maps were computed with data of the General and Violation

conditions, each time introducing an increasing percentage of randomly generated foci in the analyses, at the same threshold. Since no significant cluster emerged from the condition of Prediction Encoding at voxel-level FWE $p < 0.05$, we did not perform the fail-safe analysis in this case. The results of this procedure are generally considered robust if the clusters are still present after adding more than 200% of random data.

5.5 Leave-N-out

The Leave-N-out method is a cross-validation method to test the heterogeneity of a set of data. In the present study, it was employed on the dataset of the Prediction Encoding condition, to check if the reason behind the absence of significant clusters could be the particular impact of a study or group of studies in the meta-analysis. This method gives us the possibility to weight the contribution of each experiment (or group of experiments) and estimate the presence of "outliers", i.e. studies which contribute to the result disproportionately, driving the outcome to a certain direction [41]. This analysis has been performed by calculating the ALE results each time omitting a growing number of 'N' experiments with reinsertion after each run. Each 'leave N-out' iteration has been repeated 10 times to calculate the standard error. The calculation of 1 - quadratic error divided by the total number of experiments that are removed from the analysis and reinserted at each step is a measure called "energy". This value can be interpreted as a measure of how much the ALE results are affected by the removal of the articles, indicating how much the dataset is homogeneous. In our dataset, the level of energy that seemed to be associated with a change in its distribution is 0.6, thus the procedure at each Leave-N stopped when this threshold was reached.

5.6 Overlap between the unthresholded ALE map and the SVC Consensus map

In order to test the overall robustness of the network, obtained with the SVC Consensus technique, we created the overlap of the General Prediction Consensus and the unthresholded General ALE maps. An overlap between these maps shows brain regions that are functionally connected to each other. We tested the degree of overlap both by visually analysing the map and by calculating the Cosine Similarity Index (SIM). This is a widely used metric to assess the similarity between two vectors, which is unsensitive to their magnitude. It is calculated as the dot product of two vectors (in our case, the two maps) divided by the product of the two vectors' magnitudes. Then, the similarity between the two maps was calculated with the following formula:

$$SIM_{AB} = \frac{\sum_{k=1}^{n} A_i x\ B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\ x\ \sqrt{\sum_{i=1}^{n} B_i^2}}$$

Where A_i and B_i represent the two maps vector, and *n* the number of voxels. This index ranges from 0 to 1, where 1 indicates, in our case, a complete similarity.

Declarations

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors contributions:

L.F. conceptualized the research idea, designed the study, curated the dataset, performed part of the statistical analyses and created data visualizations. L.M. contributed to the conceptualization of this research work, provided methodological knowledge, and performed statistical analyses. D.L. devised the Supplementary Material and provided feedback on study design and dataset building. J.M., T.C. and A.T. curated substantial part of the methodology, by supporting, performing and validating the present statistical analyses. G. Z. K. provided research materials, contributed valuable input for study conceptualization, and revised the manuscript. S.D and F.C. provided organizational and methodological support, revised and corrected the manuscript.

The manuscript was primarily written by L.F., L.M. and G.Z.K., and it was approved by all the authors.

Competing interests

The authors declare neither financial nor competing interest.

References

1. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
2. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–38 (2010).
3. Huang, Y. & Rao, R. P. N. Predictive coding. *Wiley Interdiscip. Rev. Cogn. Sci.* **2**, 580–593 (2011).
4. Knill, D. C. & Pouget, A. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
5. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
6. Rauss, K., Schwartz, S. & Pourtois, G. Top-down effects on early visual processing in humans: A predictive coding framework. *Neurosci. Biobehav. Rev.* **35**, 1237–1253 (2011).

7. Osterhout, L. E. E. & Holcomb, P. J. Journal of memory and language 2002. *Lang. Cogn. Process.* **8**, 439–483 (1993).
8. Jessup, R. K., Busemeyer, J. R. & Brown, J. W. Error effects in anterior cingulate cortex reverse when error likelihood is high. *J. Neurosci.* **30**, 3467–3472 (2010).
9. Shain, C., Blank, I. A., van Schijndel, M., Schuler, W. & Fedorenko, E. fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia* **138**, 107307 (2020).
10. Gordon, N., Koenig-Robert, R., Tsuchiya, N., Van Boxtel, J. J. A. & Hohwy, J. Neural markers of predictive coding under perceptual uncertainty revealed with hierarchical frequency tagging. *Elife* **6**, 1–17 (2017).
11. Stefanics, G. & Czigler, I. Automatic prediction error responses to hands with unexpected laterality: An electrophysiological study. *Neuroimage* **63**, 253–261 (2012).
12. Wacongne, C., Changeux, J. P. & Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* **32**, 3665–3678 (2012).
13. Carbajal, G. V. & Malmierca, M. S. The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends Hear.* **22**, 1–33 (2018).
14. Cretu, A. L., Ruddy, K., Germann, M. & Wenderoth, N. Uncertainty in contextual and kinematic cues jointly modulates motor resonance in primary motor cortex. *J. Neurophysiol.* **121**, 1451–1464 (2019).
15. Hosoya, T., Baccus, S. A. & Meister, M. Dynamic predictive coding by the retina. *Nature* **436**, 71–77 (2005).
16. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**, 43–56 (2011).
17. Hohwy, J., Roepstorff, A. & Friston, K. Predictive coding explains binocular rivalry: An epistemological review. *Cognition* **108**, 687–701 (2008).
18. D'Astolfo, L. & Rief, W. Learning about expectation violation from prediction error paradigms - A meta-analysis on brain processes following a prediction error. *Front. Psychol.* **8**, 1–11 (2017).
19. Malekshahi, R. *et al.* Differential neural mechanisms for early and late prediction error detection. *Nat. Publ. Gr.* 1–13 (2016) doi:10.1038/srep24350.
20. RP, R. & DH, B. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2 VN-re**, 79–87 (1999).
21. Shipp, S. Neural elements for predictive coding. *Front. Psychol.* **7**, 1–21 (2016).
22. Bastos, A. M. *et al.* Canonical Microcircuits for Predictive Coding. *Neuron* **76**, 695–711 (2012).
23. Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P. & Pezzulo, G. Active Inference: A Process Theory. *Neural Comput.* **29**, 1–49 (2017).
24. Owens, A. P., Allen, M., Ondobaka, S. & Friston, K. J. Interoceptive inference: From computational neuroscience to clinic. *Neurosci. Biobehav. Rev.* **90**, 174–183 (2018).

25. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron* **100**, 424–435 (2018).
26. Grotheer, M. & Kovacs, G. Z. Can predictive coding explain repetition suppression? *Cortex* **80**, 113–124 (2016).
27. Kilner, J. M., Friston, K. J. & Frith, C. D. Predictive coding: An account of the mirror neuron system. *Cogn. Process.* **8**, 159–166 (2007).
28. Seth, A. K., Suzuki, K., Critchley, H. D., Frith, C. & Trust, W. An interoceptive predictive coding model of conscious presence. **2**, 1–16 (2012).
29. Raichle, M. E. A paradigm shift in functional brain imaging. *J. Neurosci.* **29**, 12729–12734 (2009).
30. Boes, A. D. *et al.* Network localization of neurological symptoms from focal brain lesions. *Brain* **138**, 3061–3075 (2015).
31. Siman-Tov, T. *et al.* Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. *Neurosci. Biobehav. Rev.* (2019) doi:10.1016/j.neubiorev.2019.08.012.
32. Chase, H. W., Kumar, P., Eickhoff, S. B. & Dombrovski, A. Y. Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cogn. Affect. Behav. Neurosci.* **15**, 435–459 (2015).
33. Summerfield, C., Monti, J. M., Trittschuh, E. H., Mesulam, M. & Egner, T. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* **11**, 1004–1006 (2008).
34. Feuerriegel, D., Vogels, R. & Kovacs, G. Z. (Friedrich-S. U. of J. Evaluating the evidence for expectation suppression in the visual system. *preprint* (2020).
35. Eickhoff, S. B. *et al.* Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Hum. Brain Mapp.* **30**, 2907–2926 (2009).
36. Turkeltaub, P. E., Eden, G. F., Jones, K. M. & Zeffiro, T. A. Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage* **16**, 765–780 (2002).
37. Lancaster, J. L. *et al.* Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum. Brain Mapp.* **28**, 1194–1205 (2007).
38. Darby, R. R., Joutsa, J. & Fox, M. D. Network localization of heterogeneous neuroimaging findings. *Brain* **142**, 70–79 (2019).
39. Sepulcre, J. *et al.* The organization of local and distant functional connectivity in the human brain. *PLoS Comput. Biol.* **6**, 1–15 (2010).
40. Acar, F., Seurinck, R., Eickhoff, S. B. & Moerkerke, B. Assessing robustness against potential publication bias in Activation Likelihood Estimation (ALE) meta-analyses for fMRI. *PLoS One* **13**, 1–23 (2018).
41. Gee, T. Capturing study influence: The concept of 'gravity' in meta-analysis. *Couns. Psychother. Heal.* **1**, 52–75 (2005).

42. Laird, A. R. *et al.* Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. *Neuroimage* **51**, 677–683 (2010).
43. Parr, T. & Friston, K. J. The anatomy of inference: Generative models and brain structure. *Front. Comput. Neurosci.* **12**, (2018).
44. Egner, T., Monti, J. M. & Summerfield, C. Expectation and surprise determine neural population responses in the ventral visual stream. *J. Neurosci.* **30**, 16601–16608 (2010).
45. Boyden, E. S., Katoh, A. & Raymond, J. L. Cerebellum-Dependent Learning: The Role of Multiple Plasticity Mechanisms. *Annu. Rev. Neurosci.* **27**, 581–609 (2004).
46. Deluca, C. *et al.* The cerebellum and visual perceptual learning: evidence from a motion extrapolation task. *Cortex* **58**, 52–71 (2014).
47. Schultz, W. & Dickinson, A. Neuronal Coding of Prediction Errors. *Annu. Rev. Neurosci.* **23**, 473–500 (2000).
48. Johnson, J. F., Belyk, M., Schwartze, M., Pinheiro, A. P. & Kotz, S. A. The role of the cerebellum in adaptation: ALE meta-analyses on sensory feedback error. *Hum. Brain Mapp.* 1–16 (2019) doi:10.1002/hbm.24681.
49. Bossaerts, P. Risk and risk prediction error signals in anterior insula. *Brain Struct. Funct.* **214**, 645–653 (2010).
50. Ishida, H., Suzuki, K. & Grandi, L. C. Predictive coding accounts of shared representations in parieto-insular networks. *Neuropsychologia* **70**, 442–454 (2015).
51. Garrison, J., Erdeniz, B. & Done, J. Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neurosci. Biobehav. Rev.* **37**, 1297–1310 (2013).
52. Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J. & Schultz, W. Neural Correlates of Value, Risk, and Risk Aversion Contributing to Decision Making under Risk. *J. Neurosci.* **29**, 12574–12583 (2009).
53. Sherman, M. T., Seth, A. K. & Kanai, R. Predictions shape confidence in right inferior frontal gyrus. *J. Neurosci.* **36**, 10323–10336 (2016).
54. Weilnhammer, V., Stuke, H., Hesselmann, G., Sterzer, P. & Schmack, K. A predictive coding account of bistable perception - a model-based fMRI study. *PLoS Comput. Biol.* **13**, 1–21 (2017).
55. DeSerisy, M., Musial, A., Comer, J. S. & Roy, A. K. Functional connectivity of the anterior insula associated with intolerance of uncertainty in youth. *Cogn. Affect. Behav. Neurosci.* **20**, 493–502 (2020).
56. Langner, R. *et al.* Modality-Specific Perceptual Expectations Selectively Modulate Baseline Activity in Auditory , Somatosensory , and Visual Cortices. (2011) doi:10.1093/cercor/bhr083.
57. Jin, H. *et al.* Involvement of the left inferior frontal gyrus in predictive inference making. *Int. J. Psychophysiol.* **71**, 142–148 (2009).
58. Heilbron, M. & Chait, M. Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience* **389**, 54–73 (2018).

59. Vassena, E., Krebs, R. M., Silvetti, M., Fias, W. & Verguts, T. Neuropsychologia Dissociating contributions of ACC and vmPFC in reward prediction , outcome , and choice. *Neuropsychologia* **59**, 112–123 (2014).
60. Uddin, L. Q. Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* **16**, 55–61 (2015).
61. Cavanna, A. E. & Trimble, M. R. The precuneus: A review of its functional anatomy and behavioural correlates. *Brain* **129**, 564–583 (2006).
62. Strauss, M. *et al.* Disruption of hierarchical predictive coding during sleep. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1353-62 (2015).
63. Mumford, D. On the computational architecture of neocortex. *Biol. Cybern.* **65**, 135–145 (1991).
64. Carter, C. S. *et al.* Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science (80-.).* **280**, 747–749 (1998).
65. Critchley, H. D., Tang, J., Glaser, D., Butterworth, B. & Dolan, R. J. Anterior cingulate activity during error and autonomic response. *Neuroimage* **27**, 885–895 (2005).
66. Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9673–9678 (2005).
67. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
68. Fox, M., Corbetta, M., Snyder, A., Vincent, J. & Raichle, M. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10046–51 (2006).
69. Schroeder, C. E. & Lakatos, P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* **32**, 9–18 (2009).
70. Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. & Schroeder, C. E. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science (80-.).* **320**, 110–113 (2008).
71. Busch, N. A. & VanRullen, R. Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci.* **107**, 16048–16053 (2010).
72. Schröger, E., Marzecová, A. & Sanmiguel, I. Attention and prediction in human audition: A lesson from cognitive psychophysiology. *Eur. J. Neurosci.* **41**, 641–664 (2015).
73. Hsu, Y.-F., Hamalainen, J. & Waszak, F. Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Front. Hum. Neurosci.* **8**, 152 (2014).
74. Aitchison, L. & Lengyel, M. With or without you: predictive coding and Bayesian inference in the brain. *Curr. Opin. Neurobiol.* **46**, 219–227 (2017).
75. Nahum, L., Barcellona-Lehmann, S., Morand, S., Sander, D. & Schnider, A. Intrinsic emotional relevance of outcomes and prediction error: Their influence on early processing of subsequent stimulus during reversal learning. *J. Psychophysiol.* **26**, 42–50 (2012).

76. Ransom, M., Fazelpour, S. & Mole, C. Attention in the predictive mind. *Conscious. Cogn.* **47**, 99–112 (2017).
77. Summerfield, C. & Egner, T. Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* **13**, 403–409 (2009).
78. Clark, A. *Surfing uncertainty: Prediction, action, and the embodied mind*. (Oxford University Press, 2015).
79. Friston, K. & Feldman, H. Attention , uncertainty and free-energy. *Hum. Neurosci.* **4**, 1–23 (2010).
80. Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C. & De Lange, F. P. Attention reverses the effect of prediction in silencing sensory signals. *Cereb. Cortex* **22**, 2197–2206 (2012).
81. Chennu, S. *et al.* Expectation and attention in hierarchical auditory prediction. *J. Neurosci.* **33**, 11194–11205 (2013).
82. Thillay, A. *et al.* Sustained attention and prediction: distinct brain maturation trajectories during adolescence. *Front. Hum. Neurosci.* **9**, (2015).
83. Andermane, N., Bosten, J. M., Seth, A. K. & Ward, J. Individual differences in the tendency to see the expected. *Conscious. Cogn.* **85**, 102989 (2020).
84. Jones, A., Hughes, G. & Waszak, F. The interaction between attention and motor prediction. An ERP study. *Neuroimage* **83**, 533 (2013).
85. Marzecová, A., Widmann, A., SanMiguel, I., Kotz, S. A. & Schröger, E. Interrelation of attention and prediction in visual processing: Effects of task-relevance and stimulus probability. *Biol. Psychol.* **125**, 76–90 (2017).
86. Smout, C. A., Tang, M. F., Garrido, M. I. & Mattingley, J. B. Attention promotes the neural encoding of prediction errors. *PLoS Biol.* **17**, 1–22 (2019).
87. Fox, K. C. R., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R. & Christoff, K. The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage* **111**, 611–621 (2015).
88. Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. The brain's default network: Anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* **1124**, 1–38 (2008).
89. Klarić, K. The World According To My Predictions: Human Brains' Default Mode Network In The Context of Predictive Coding. 8 (2018) doi:10.13140/RG.2.2.23541.01767.
90. Carvalho, F. M., Chaim, K. T., Sanchez, T. A. & de Araujo, D. B. Time-Perception Network and Default Mode Network Are Associated with Temporal Prediction in a Periodic Motion Task. *Front. Hum. Neurosci.* **10**, 268 (2016).
91. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12574–12579 (2016).
92. Tomasi, D. & Volkow, N. D. Functional connectivity hubs in the human brain. *Neuroimage* **57**, 908–917 (2011).

93. Vatansever, D., Menon, D. K. & Stamatakis, E. A. Default mode contributions to automated information processing. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12821–12826 (2017).
94. Pezzulo, G., Zorzi, M. & Corbetta, M. *The secret life of predictive brains: what's spontaneous activity for?* vol. in press (2020).
95. Roepstorff, A. Interactively human: Sharing time, constructing materiality. *Behav. Brain Sci.* **36**, 224–225 (2013).
96. Feuerriegel, D. Selecting appropriate designs and comparison conditions in repetition paradigms. *Cortex* **80**, 196–205 (2016).

Figures

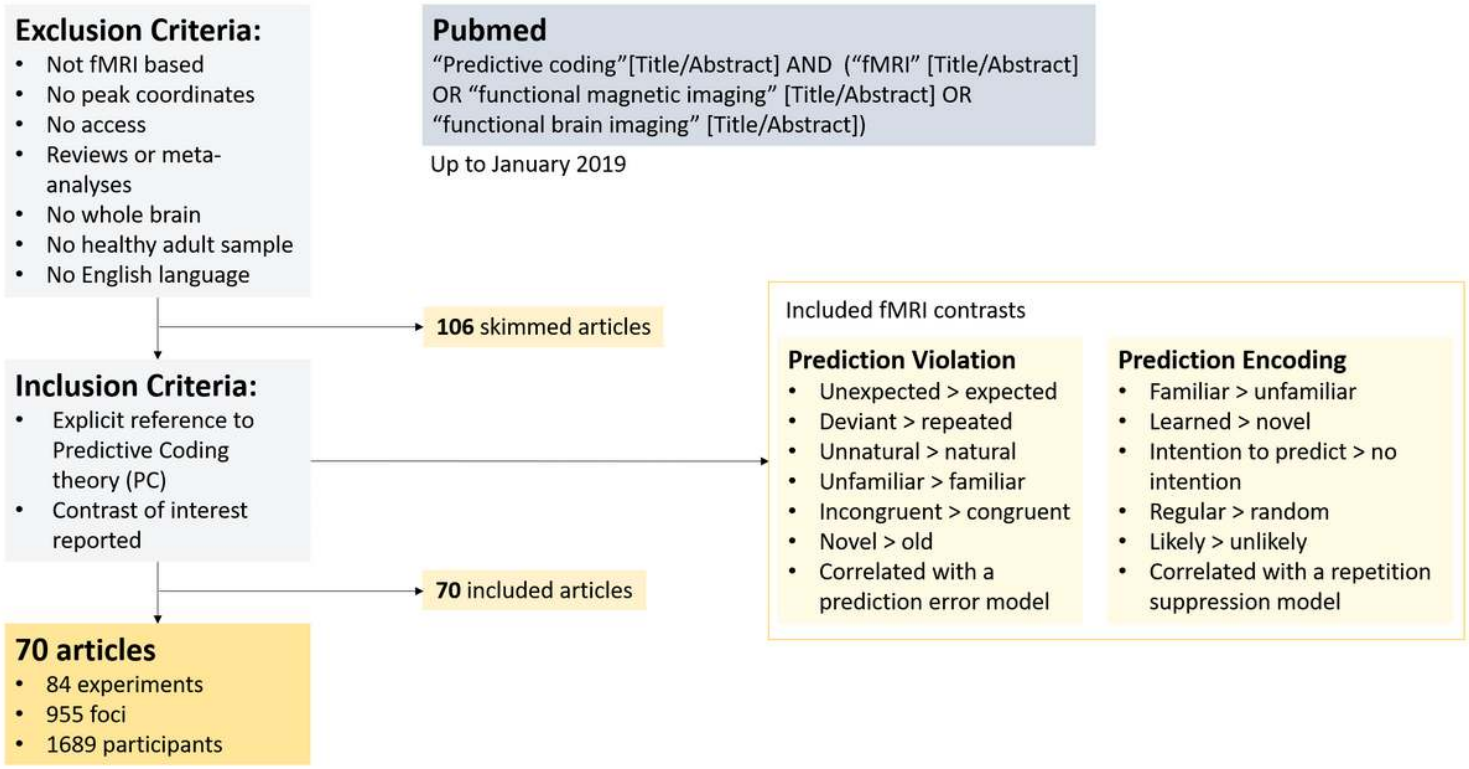

Figure 1

Flowchart representing the process of search and selection of the eligible articles for the meta-analysis and the SVC Consensus.

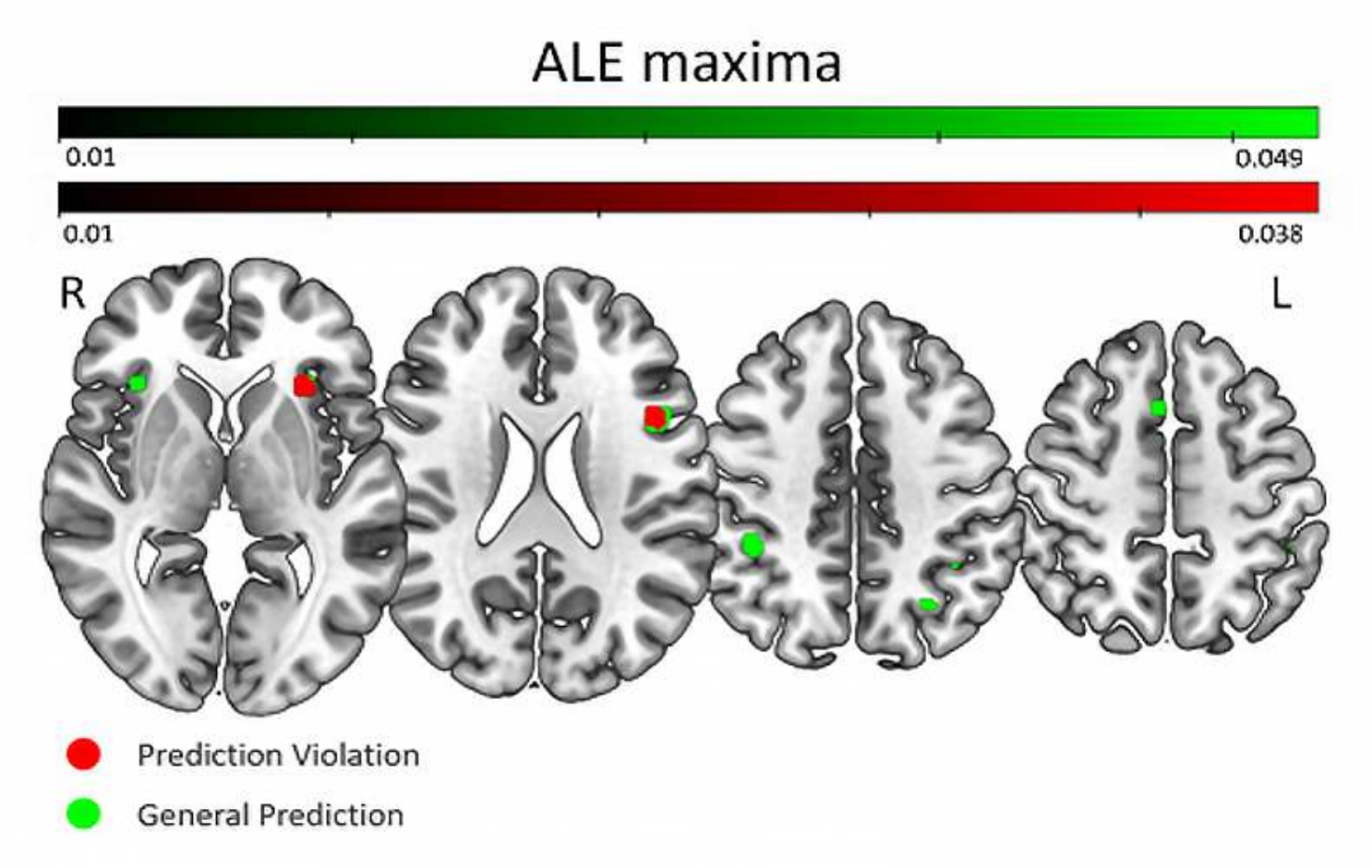

Figure 2

Activation Likelihood Estimation results at a FWE corrected voxel-level threshold ($p<0.05$). Green: condition of General Prediction; Red: condition of Prediction Violation. Two clusters are in common between the two conditions, one in the left anterior insula/claustrum and the other in the left inferior frontal gyrus/precentral gyrus. The General Prediction condition also shows clusters in the right insula, right and left inferior parietal lobule, one in the cuneus and one in the right middle frontal gyrus.

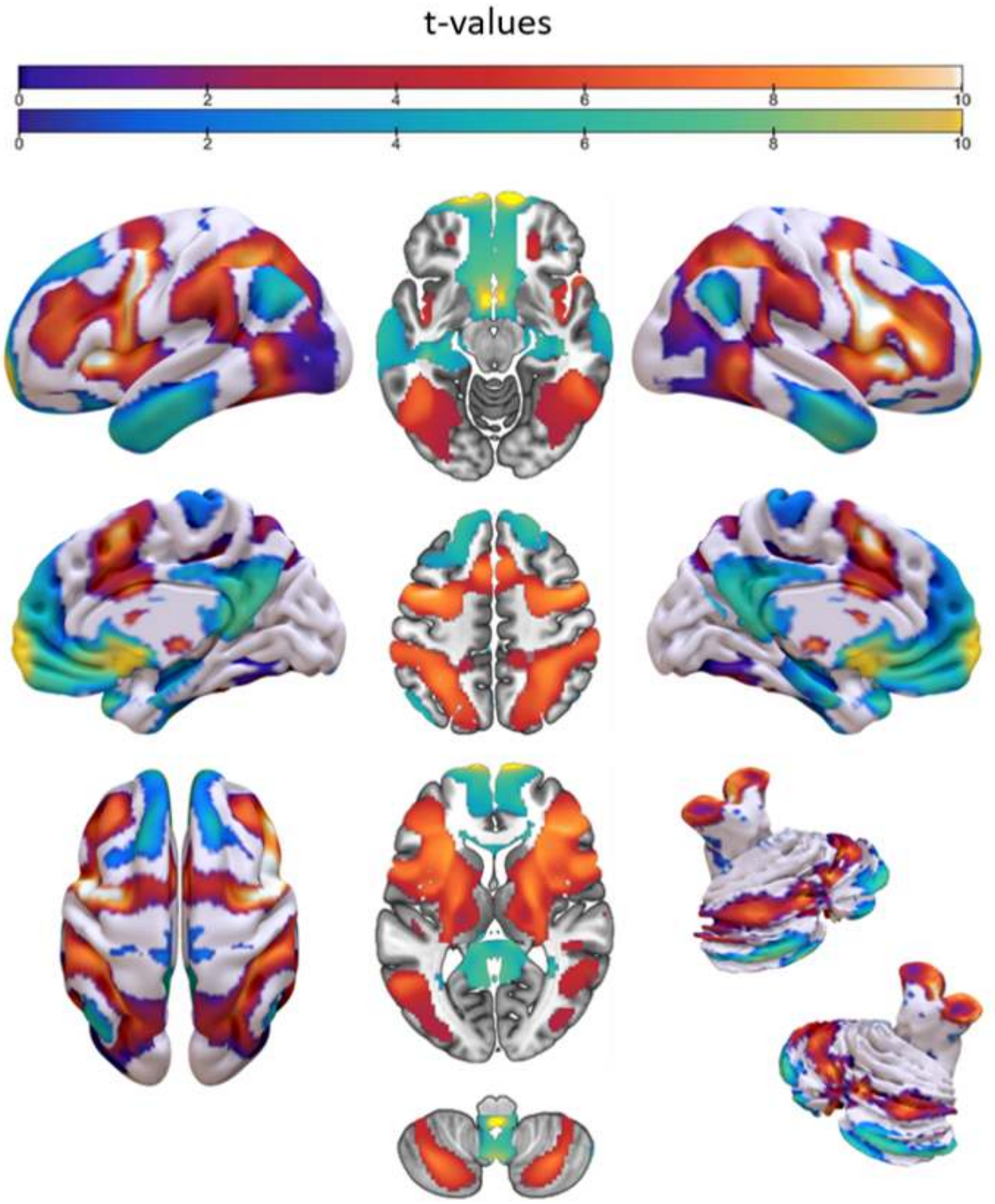

Figure 3

Surface, medial and cerebellar mapping of the SVC consensus analysis revealing the “predictive network” for the condition of General Prediction. Note that this and the Prediction Violation areas overlap entirely. The network that shows the Prediction Encoding areas is presented in the Supplementary Materials (Figure S2). Warm colors represent positive t-values (range: 4.37-14.17), cold colors are for negative t-values (range: 4.37-19.61), shown in arbitrary units.

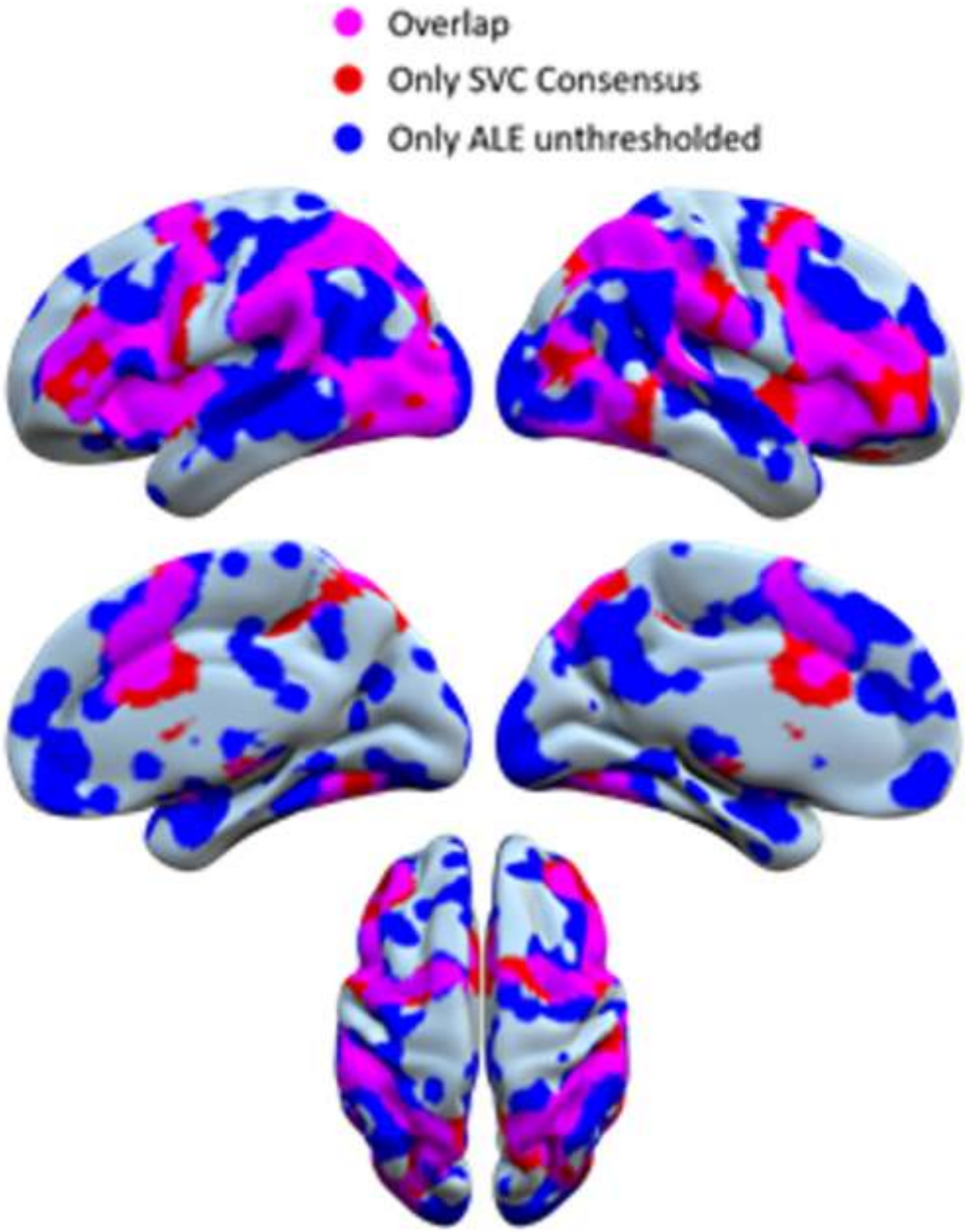

Figure 4

Surface map showing the overlap between the results of the unthresholded ALE (blue) and the SVC Consensus (red) analyses, for General Prediction. The overlapping regions are presented in purple

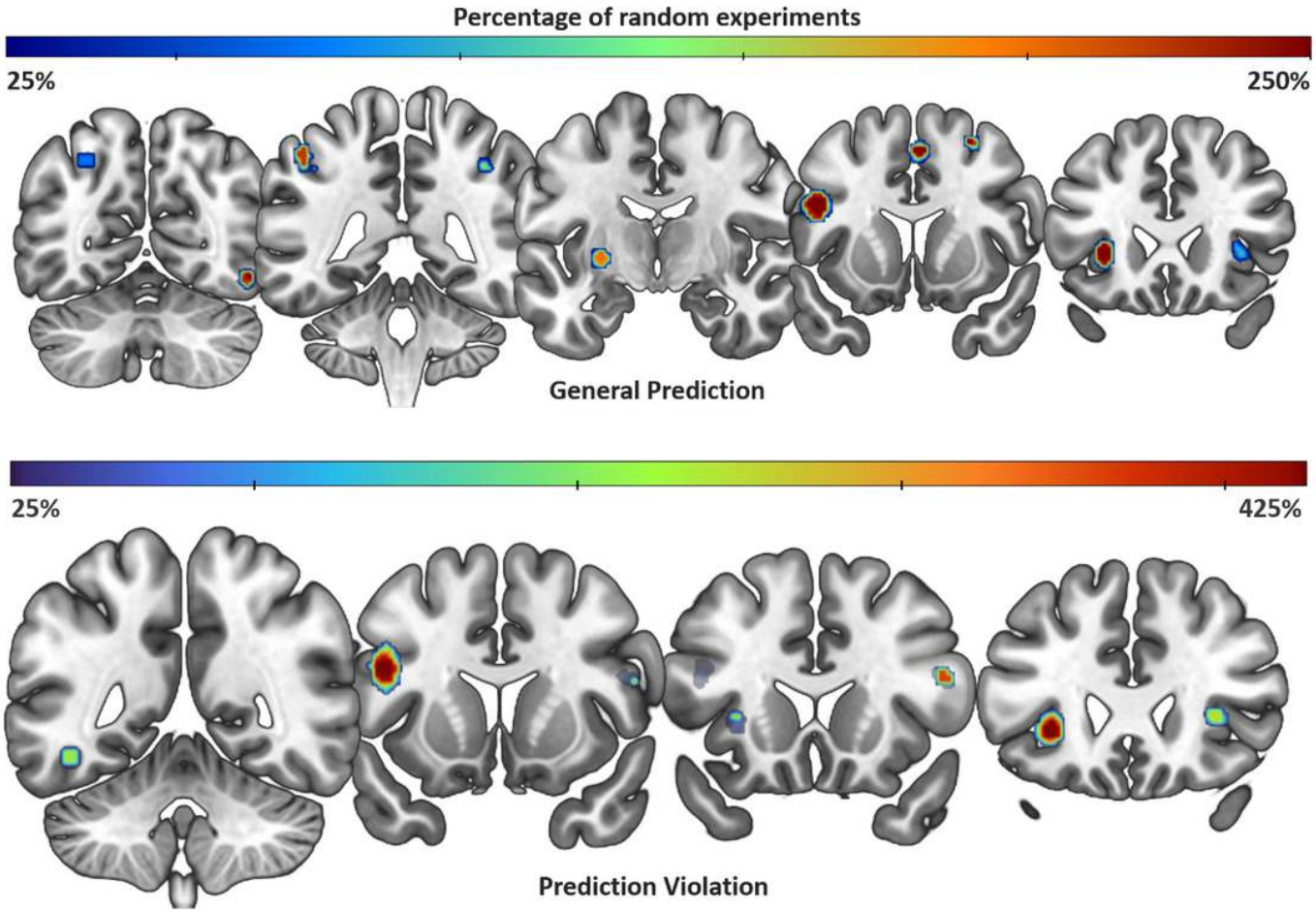

Figure 5

Coronal sections showing the results of the fail-safe analysis. Upper row: General Prediction; lower row: Prediction Violation. The color scales represent gradually increasing added random data in percentage. The warmer the color of the pixels, the more random noise is tolerated by the voxels.

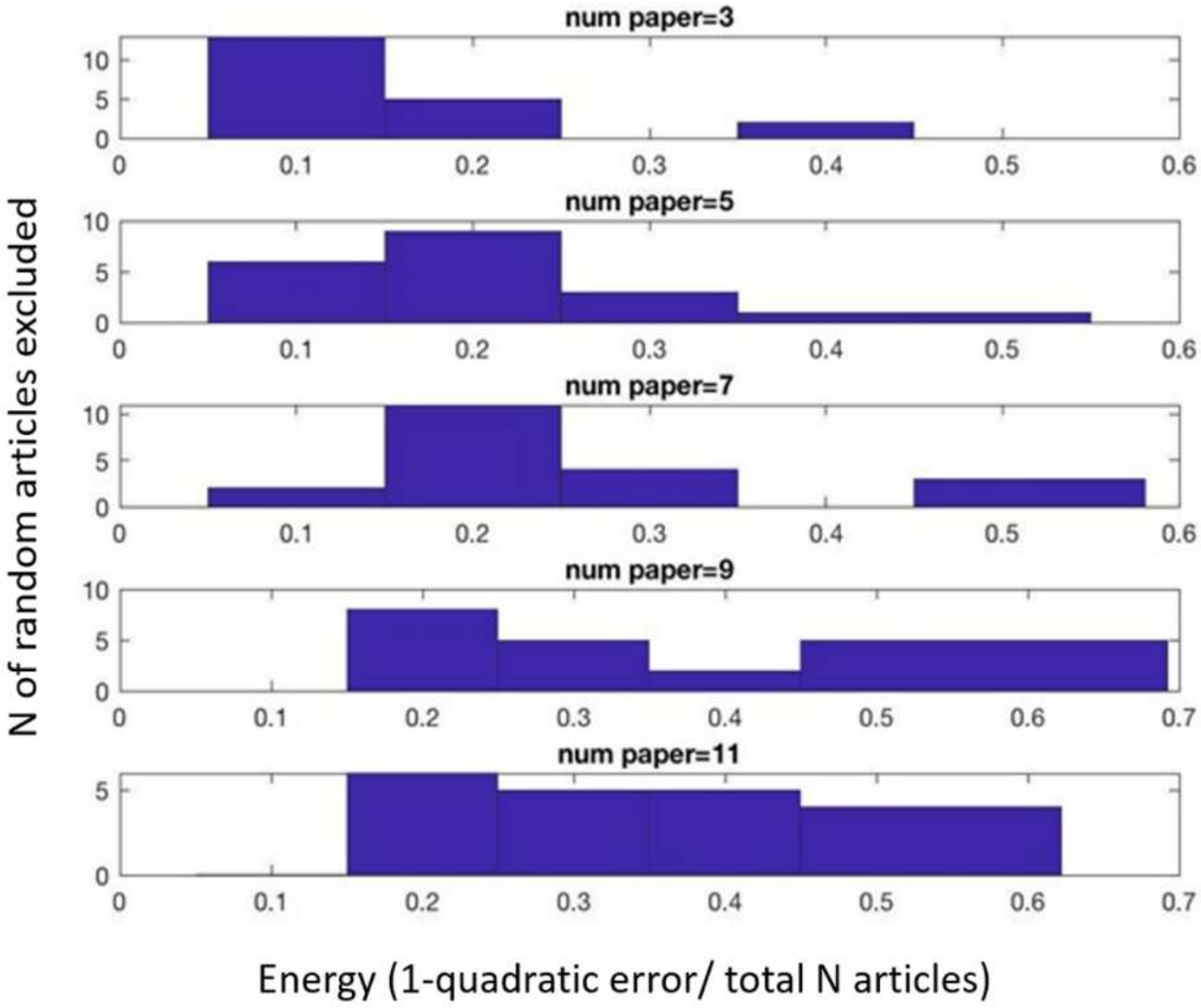

Figure 6

Diagram relative to the results of the Leave-N-Out analysis on the articles included in the Prediction Encoding condition. In each section, the y-axis shows the number of articles, and the x-axis shows the energy (1-quadratic error/total N experiments). To calculate the standard error, each run with removal and reinsertion of 'N' articles was repeated 10 times. Since removing up to 7 articles at a time (18% or the total dataset) still did not introduce important changes in the distribution of energy, we conclude that the dataset is homogeneous. See Section 5.6 for more details.

Supplementary Files

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

- FiccoPredictiveNetworkSupplementaryInformation.pdf