Evidence for embodied predictive coding: the anterior insula coordinates cortical processing of tactile deviancy

Micah Allen[1,2], Francesca Fardo[3], Martin J Dietz[3], Hauke Hillebrandt[1,4], Geraint Rees[1,2], Andreas Roepstorff[3,5]

1. Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom,
2. Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom, and
3. Center of Functionally Integrative Neuroscience, Aarhus University Hospital, Aarhus 8000, Denmark
4. Harvard University, Cambridge, MA, 02138, United States
5. Interacting Minds Centre, Aarhus University, DK-8000 Aarhus C, Denmark

ABSTRACT

Embodied awareness is the pervasive, multimodal self-awareness that is thought to form the foundation of emotion. This awareness was recently proposed to rely on the anterior insular cortex (AIC) comparing expected and actual bodily signals arising in prefrontal and sensory cortices. To investigate this possibility in the somatosensory domain, we measured brain activity using functional magnetic resonance imaging while healthy participants discriminated tactile stimuli in a roving oddball design. Dynamic Causal Modelling revealed that unexpected stimuli increased the strength of forward connections in a caudal to rostral ascending hierarchy from thalamic and somatosensory regions towards insula, cingulate and prefrontal cortices, consistent with hierarchical predictive coding. Within this feed-forward flow of neural coupling, the AIC increased both forwards and backwards connections with prefrontal and somatosensory cortex, supporting a comparator role. Further, we found that greater prefrontal to AIC connectivity predicted subjective ratings of stimulus discrimination difficulty. These results are interpreted in light of embodied predictive coding, suggesting that the AIC coordinates global cortical processing of tactile changes to support body awareness.

INTRODUCTION

The sense of self and experience of emotion are known to rely on highly multimodal interactions between interoceptive, somatosensory and proprioceptive signals, as well as prior beliefs about the causes of these sensations (Gallagher 2005; Brown et al. 2013). Recently, predictive coding has been used to explain how these interactions produce embodied and emotional awareness (Apps and Tsakiris 2013; Seth 2013). One key prediction is that the anterior insula cortex (AIC) acts as a comparator between incoming bodily sensations and the brain's prior beliefs. In this paper we test this hypothesis in the somatosensory domain, using dynamic causal modelling (DCM).

Somatosensation underlies the conscious body-image that underpins body-ownership and action agency (Serino and Haggard 2010) and has recently been shown to rely on hierarchical Bayesian inferences (Auksztulewicz et al. 2012; Ostwald et al. 2012). The AIC is richly interconnected with the posterior insula and somatosensory cortex (Cerliani et al. 2012; Chang et al. 2012), and anticipates the sensory and affective consequences of touch (Lovero et al. 2009). Probabilistic hierarchical inferences involving the AIC and cingulate cortex are thought to support interoception (Seth et al. 2011), pain analgesia (Büchel et al. 2014), and somatosensation (Ostwald et al. 2012). The AIC in particular is thought to integrate these modalities to create a multimodal self-model (Apps and Tsakiris 2013; Gu et al. 2013) that underpins dynamic cognitive control (Ullsperger et al. 2010) and embodied self-awareness (Gallagher 2005). Further evidence comes from findings linking AIC activations to body-state prediction errors, for example in the case of gustatory conflict (O'Doherty et al. 2002), unexpected painful stimulation (Seymour et al. 2005; Keltner et al. 2006), or during false heart rate-feedback (Critchley et al. 2004). Little is currently known about how the AIC communicates tactile prediction errors throughout the cortical hierarchy. Delineating the impact of somatosensory changes on insula responses is thus critical for understanding embodied perception.

To address these issues, we adapted a roving somatosensory oddball task (RSOT) for use during fMRI scanning. Sensory oddball or deviance responses in the brain have been strongly linked to prediction error minimization, and the RSOT has previously been shown to elicit prediction errors in somatosensory and cingulate cortex. An advantage of the RSOT over standard oddball designs is the ability to control for stimulus-specific effects, as all stimuli (e.g. high and low intensity) serves as both deviant and standard. We thus used DCM for fMRI to investigate the causal impact of unexpected somatosensory stimuli on effective neural connectivity between the AIC and other structures. We reasoned that if the AIC acts as a bodily comparator, deviance should increase both ingoing connections from sensory and thalamic areas and outgoing connections to prefrontal and cingulate cortex. We further expected to observe an overall pattern of increased forward-driving connectivity throughout a hierarchy of subcortical and cortical regions in response to deviancy, consistent with a hierarchical predictive coding account of somatosensory awareness. Finally, if top-down expectations underlie bodily awareness (Seth et al. 2011; Apps and Tsakiris 2013), we predicted that participants with the strongest modulation of backwards connections from the prefrontal cortex to the AIC would rate somatosensory deviants as easier to detect.

METHOD

Participants

Thirty-eight healthy participants (16 males) were recruited from Aarhus University and the surrounding community. Inclusion criteria specified that all participants were between the ages of 21-45 years, right handed, free from medications with contraindications for the BOLD signal (psychiatric, blood pressure or heart medication, etc.), physically and mentally healthy, and meeting standard MRI safety inclusion criteria (lack of claustrophobia, metallic implants, etc.). All participants gave verbal consent and visited the MRI laboratory at Aarhus University Hospital for approximately 2 hours in total, and received a 300 DKK (approx. €40) participation reimbursement. All experimental procedures were conducted with approval from the local ethics committee (De Videnskabsetiske Komitéer for Region Midtjylland) in accordance with the Declaration of Helsinki. 8 participants in total were excluded from preliminary data analysis - one for excessive motion during scanning, 6 for extremely poor behavioral performance (see Roving Somatosensory Oddball Task for further details), and one for failure to acquire pulse regressors. The final sample for the fMRI analysis included 30 participants (14 males) with a mean age of 24.5 years (SD = 3.2).

Roving Somatosensory Oddball Task

To manipulate tactile probability while controlling for stimulus intensity and attention, we utilized a Roving Somatosensory Oddball Task (RSOT) in which trains of stimuli randomly switch between high and low intensity after a variable number of repetitions (Garrido et al. 2008; Ostwald et al. 2012). In the present study, stimuli were delivered in trains of 3-7 repetitions. Stimuli consisted of single electrical pulses of 50 µs duration and 2000 ms interstimulus interval. Following each repetitive train, stimuli switched between low or high intensity trials, where low intensity trials corresponded to a single pulse at twice the perceptual threshold, and high intensity trials consisted of two pulses identical to the single delivered in rapid succession (100 ms inter-stimulus interval). This paradigm has been repeatedly shown to elicit sensory prediction errors (Garrido et al. 2008, 2009; Lieder et al. 2013), with deviant trials selectively increasing the strength of forward-driving neural connectivity (Dietz et al. 2014) and eliciting Bayesian surprise in somatosensory and cingulate cortices (Ostwald et al. 2012).

This stimulation protocol resulted in a sensation of a mild tickle or vibration that was not reported as painful by any participant. The first stimulus of each new train was modelled as the “deviant”, and the third repetition following the deviant as the “standard”. The number of repetitions between each switch was randomly sampled from a normal distribution over possible repetition numbers (i.e., between 3 and 7), generating an unpredictable uniform stimulus sequence. Participants received a total of 158 deviant and 640 repetition stimuli (of which 158 stimuli were selected as standard). All stimuli were delivered to the median nerve of the left forearm using two MR-safe ECG electrodes placed approximately 2.5 cm apart and a constant current stimulator (DeMeTec, Langgoens, Germany). See Figure 1 for an overview of the experimental set-up and sample stimulus train.

After placement in the scanner, participants’ individual perceptual thresholds were determined using an adaptive staircase procedure prior to scanning. The staircase consisted of a one-up/three-down procedure, where step size was reduced every two reversals until reaching a total of 8 reversals. The sensory threshold was thus calculated by averaging the stimulus intensities corresponding to the 8 reversals. Stimuli for the subsequent oddball task were then delivered at twice this sensory threshold, eliciting a mild touch sensation. After thresholding, participants completed a short practice version of the oddball task, and continued to the main experiment after indicating that the task instructions were fully understood. All participants completed approximately 30 minutes of the RSOT during fMRI acquisition.

Pilot investigation with this stimulation protocol revealed that as intended, the double stimulation was perceived as slightly more intense than single trials. To control attention, participants were instructed to silently count all stimuli switches throughout the entire task duration, in a standard ‘active’ counting task (Garrido et al. 2009). This manipulation ensures that participants must exert equivalent attentional effort to both deviant and standard trials, as the occurrence of deviants is unpredictable. Participant switch counts were then recorded at the end of the imaging session to ensure compliance. Six participants reporting switch counts 60% above or below the true total (i.e., poorer than chance performance) were excluded from data analyses. Overall switch count accuracy of the remaining participants was extremely high (mean accuracy 99%), suggesting successful attentional control and task participation.

Following the scan, participants completed a debriefing inquiring about the nature of the felt stimuli (e.g., painful or non-painful). Participants also rated the felt intensity of each stimulus type (i.e., low and high), and the difficulty detecting stimulus changes from low to high and from high to low, on visual analog scales with 0 marked as ‘not at all intense/difficult’ and 100 labeled as ‘very difficult/intense’. The adaptive staircase procedure, the RSOT and the post-scan ratings were implemented in Psychopy (v1.76.00) (Peirce 2007).

Data acquisition and preprocessing

All brain measurements were acquired on a Siemens Trio 3T scanner, using a 32-channel head coil. For fMRI, 31 slices were acquired in ascending order using a gradient echo planar sequence with echo time 30 ms, voxel size 3 × 3 × 3 mm in a 64 × 64 mm field of view, repetition time = 1.54 s. Slices were manually positioned to ensure full coverage of somatosensory cortex, anterior insula, prefrontal cortex, and thalamus, flip angle = 90°. A T1-weighted MPRAGE structural image was collected after the EPI sequence. B_0 field maps were collected using a gradient echo field map sequence. To control for physiological BOLD-signal confounds, cardiac cycles were recorded in synchrony with EPI acquisition using an infrared pulse oximeter on the participant’s right index finger.

fMRI Analysis

MRI data were analysed using Statistical Parametric Mapping (SPM8 for GLM analysis and SPM12b for DCM, http://www.fil.ion.ucl.ac.uk/spm). Each participant's 1,109 EPI images were corrected for geometric distortions caused by susceptibility-induced field inhomogeneities. This was done using a combined correction for both static distortions and changes in those distortions caused by head motion (Andersson et al. 2001; Hutton et al. 2002). Static distortions were calculated using the FieldMap toolbox to process each participant's B_0 field map (Hutton et al. 2004). EPI images were then realigned, unwarped, and co-registered to the participant's anatomical scan. The anatomical images were processed using the unified segmentation procedure implementing tissue segmentation, bias correction, and spatial normalization (Ashburner and Friston 2005); derived normalization parameters were then applied to the EPI images. Finally, the images were smoothed using a 6 mm full-width at half-maximum Gaussian kernel, and resampled to 3 × 3 × 3 mm voxels.

To control for motion and physiological BOLD signal confounds, serial correlations were modelled using a nuisance variable regression approach thoroughly described in (Lund et al., 2006). In addition to the SPM8 standard discrete cosine set high pass filter (128 s cut off), this approach included 10 RETROICOR-derived regressors based on cardiac oscillations (Glover et al., 2000). We also included the full 12 parameter Volterra expansion of motion and motion history parameters to capture rigid body head movement related to subject motion and respiration (Friston et al., 1996).

To examine BOLD responses to somatosensory oddballs, we modelled the EPI data in a fixed-effects general linear model (GLM) for individual participant BOLD timeseries. To doso we modelled Deviants (the first trial of a new stimulus intensity) and Standards (the third repetition following each Deviant) as separate event-related regressors convolved with the canonical hemodynamic response function. The remaining repetition trials were treated as implicit baseline. Mass-univariate statistical analysis was conducted using a hierarchical t-contrasts; fixed effects within each participant were assessed using a Deviant > Standard unidirectional t-contrast. The resulting contrast images were then passed to a random-effects one sample t-test over all participants, contrasting for positive mean response. The resulting SPM was peak-corrected for multiple comparisons at a family-wise error rate $P_{FWE} < 0.05$ using Gaussian random field theory (Nichols and Hayasaka 2003).

Dynamic Causal Modelling

For our analysis of effective connectivity, we were interested in the overall impact of surprising tactile changes on effective connectivity within the right lateralized network identified by our mass-univariate analysis. Our connectivity hypotheses primarily concerned the pattern of modulations within a relatively large network of sub-cortical, cortical, and prefrontal areas rather than the intrinsic network architecture (e.g. the presence or absence of connections). To accommodate this, we utilized a recently developed data-driven approach to model selection that is optimised for large scale parameter estimates (Friston and Penny 2011; Friston et al. 2011). This allowed us to search within a "full" model, which contained all free and intrinsic parameters, for the best of all its possible sub-models. Within this best model, we applied classical parametric tests to the strength of specific connections. This approach is based on the high a priori plausibility of reciprocal mono- and polysynaptic connections (both of which are captured by directed connections in DCM for fMRI) existing between the majority of cortical and sub-cortical centres (Friston et al. 2011; Bastos et al. 2012; Markov et al. 2013) and thus prioritizes inferences over parameter strengths. We were thus able to make robust inferences about the strength and directionality of connections within an inclusive model space while circumventing limitations regarding the combinatorial explosion of models in classical DCM approaches (Lohmann et al. 2012).

To do so, we first remodelled our design to include the deviance condition as a parametric connectivity modulator, recoding all trials into a single regressor parametrically modulated by the Deviant > Standard (D>S) contrast (Stephan et al., 2010). We then extracted BOLD time series for the main effect of D>S from each volume of interest (VOI), for use during the specification of our DCMs. This extraction was based on peak activations in the Deviant > Standard group contrast, with VOIs in the dorsal-posterior thalamus (TH) [MNI_{xyz} = 12, -16, 10], somatosensory area 2 [MNI_{xyz} = 48, -34, 49] (S1), anterior insula cortex (AIC) [MNI_{xyz} = 36, 20, 1], anterior mid-cingulate cortex (MCC) [MNI_{xyz} = 3, 23, 43], and middle frontal gyrus (MFG) [MNI_{xyz} = 36, 50, 22]. All anatomical labels at extracted coordinates were confirmed using the SPM Probabilistic Anatomy Toolbox (Eickhoff et al. 2005).

VOI time series were then extracted from the D > S contrast in each participant via an automated search procedure. To do so, time series were summarized as the principle eigenvariate extracted from within a 6mm spherical VOI centred on each participant's local maxima. The position of each local maximum was determined by searching within a 12 mm radius search sphere (i.e., twice our 6mm FWHM smoothing kernel), centred on the group coordinate for that region. Extracted peak coordinates were plotted on a standard brain and visually inspected to ensure all extracted time series were from the appropriate anatomical region of interest. For extraction, participant-level SPMs were thresholded at $p < 0.05$ uncorrected, voxel extent threshold $k > 5$ contiguous voxels. All time-series were corrected for the effects-of-interest F-contrast. Given that we stimulated the left median nerve and were primarily interested in the right-lateralized body-awareness network (Craig 2003), all VOIs were extracted from the right side of the brain. In five participants, regional VOIs from one or more regions could not be obtained, leaving 25 total participants for all DCM analyses.

We chose a somewhat conservative approach to modelling the oddball effect by using the difference between deviants and standards as both a driving and modulatory input. This reflects the fact that, from the point of view of fMRI, the repeated presentation of stimuli every two seconds (as in our design) is effectively a steady state stimulus. Therefore, the only events inducing a haemodynamic response are the occasional deviants (relative to an arbitrary standard). We thus allowed for the deviant input to exert driving and modulatory effects. In other words, we allowed for a direct exogenous effect of deviants (mediated by unknown sources) and an effect mediated by a change in the sensitivity to extrinsic and intrinsic afferents from modelled sources (see Figure 3C for illustration of the full model). The full model thus included the impact of deviants as a driving input to the thalamus, and modulatory deviance effects on all intrinsic and self-connections. The full model with extrinsic (fixed) connections between all nodes and deviance vs standard modulations (free parameters) for all connections was then estimated in each participant using the variational Bayesian expectation maximization algorithm implemented in SPM12.

For model selection, we applied the post-hoc Bayesian model optimization procedure for network discovery (Friston et al. 2011). This technique estimates the evidence for all possible models by inverting the full model and applying a greedy search algorithm to find the probability of particular connections existing and whether a connection is modulated by an experimental condition (Friston and Penny 2011). The post-hoc model optimization routine thus furnished posterior model probabilities (i.e., the probability that a model is the best explanation for the data) for all reduced models, including a null-model with no connections. The strength of evidence for the winning model was determined using the Bayes factor; i.e., the ratio of evidence for the best model vs. the second best model (Penny et al. 2004). The parameter estimates from the winning model for each participant were then subjected to conventional frequentist analyses to determine relative modulatory strengths for various connections and contrasts of connections. These tests were false discovery rate corrected at $P_{FDR} < 0.05$ for multiple comparisons.

We tested two connectivity hypotheses; first, we established the principal directionality of deviant-driven connectivity modulations in terms of the strength with which deviant vs. standard trials modulated each of the 25 intrinsic and self-connections. This was accomplished using one-sample t-tests over the 25 estimated modulatory (B-matrix) parameters, $P_{FDR} < 0.05$. This approach enabled us to establish both the overall pattern of deviance-evoked changes in on effective connectivity within the somatosensory-oddball network, and to specifically evaluate the directionality of modulations to and from the AIC. Second, we assessed the relationship between individual differences in participants' perceived difficulty detecting sensory changes (i.e., the averaged post-scan difficulty ratings) and deviance-driven modulation of each intrinsic connection (i.e., the 20 between-region B-matrix parameters). To do so, we conducted robust regression analyses using Tukey's Biweight. This method was chosen over a least squares approach to protect against outlier values, which are a frequent issue in neuroimaging individual differences analyses (Poldrack 2012). Regression *p*-values were adjusted for multiple comparisons to a $P_{FDR} < 0.05$. All ANOVA and one-sample *t*-test analyses were conducted in SPSS version 20 (IBM), and all FDR thresholds and robust regression analysis were calculated using MATLAB R2012b (Mathworks, Inc) and the FDR toolbox.

Results

Sample Characteristics and Nuisance Regression

Minus the 6 participants excluded for extremely poor performance, the average number of counted deviants was 156 (SD = 17) out of 158 total, corresponding to an average count accuracy of 99%. This result indicates that the majority of participants were able to fully comply with the task instructions, precluding major differences in attentional effort between standard and deviant trials. In the post scan debriefing, all participants reported that the stimuli were perceived as a non-painful mild touch sensation. The average sensory threshold across participants was 12.22 mA (SD = 2.86). As a manipulation check, we compared participant's intensity and difficulty ratings for low vs. high stimuli via paired-sample t-tests. Double stimuli (mean intensity rating = 57, SD = 21) were rated as significantly more intense than single stimuli (mean intensity rating = 46, SD = 16, mean difference = 11, SD = 19, $t_{29} = 3.4$, p = 0.002), validating our stimulus manipulation. As no significant difference was found for the self-rated difficulty of discriminating single-to-double (mean difficulty rating = 34, SD = 26) or double-to-single trials (mean difficulty rating = 36, SD = 23, mean difference = -1.5, SD =17.5, $t_{29} = -0.5$, $p = .64$), we averaged the two difficulty ratings from each participant to derive an index of change awareness. This index was then used as an independent variable in our regression analysis with DCM modulatory parameters.

Mass-univariate results

As expected, our fMRI GLM analysis of the Deviant > Standard contrast revealed extensive bilateral activations in primary somatosensory and parietal cortex. Within the right hemisphere, somatosensory activations covered 27.8% of area 2 and extended into areas of the intra-parietal cortex. The largest proportion of this activation was within area 2 (7.8% of cluster) followed by the IPC (6.1%). Consistent with previous oddball fMRI studies we additionally observed significant bilateral activations in the dorsal mid-cingulate, anterior insula, and middle frontal gyrus extending into dorsolateral prefrontal cortex (BA 45). In the midbrain, we observed bilateral activations in dorsal-posterior thalamus (identified as thalamus-prefrontal using the SPM anatomy toolbox), and caudate nucleus. All anatomical labels and percent activations were determined using the SPM probabilistic anatomy toolbox. See Table 1 and Figure 2 for a complete overview of these results.

DCM Results

Post-hoc model optimization found that the full model (M255, shown in Figure 4B), with all intrinsic connections and modulations, had the highest posterior probability (pP = 0.79). The next most probable model was M128 with a posterior probability of 0.06; the Bayes factor discriminating these two models (pP_{M255}/pP_{M128}) was 13.17, corresponding to positive evidence for model 255 being the best explanation for the data (Penny et al. 2004). See Figure 3 for an overview of the model selection results and plots illustrating fixed and modulatory connectivity strengths for the winning model.

One sample t-tests over all 25 modulatory parameters revealed a general pattern of increased modulation by somatosensory deviants in a forward driving caudal to rostral hierarchy, with significant increases in connections from TH to AIC and S1, from S1 to AIC, MCC, and MFG, from AIC to MCC and MFG, and from MCC to MFG (Figure 4). In line with the hypothesis that AIC acts as a body-state comparator, the AIC increased both backwards connectivity towards S1 and forwards connectivity to the MCC and MFG. The AIC and S1 were the only regions to show increases in reciprocal connectivity. Additionally, significant modulations of the TH, AIC, and S1 self-connections were found, suggesting that somatosensory oddballs induce strong dis-inhibition of these regions. Finally, our robust regression analyses found three modulatory effects significantly predicting subjective difficulty ratings (Figure 5B), all on backwards connections (MFG to TH, MFG to MCC, and MFG to AIC). Only the MFG to AIC relationship survived FDR correction, with the Deviant > Standard modulation predicting 38% of the variance in subjective difficulty; $t(1, 25) = -3.47$, pFDR = 0.0015, $R^2 = 38.04$ (Figure 5A).

Discussion

In the present study we demonstrated that BOLD responses to surprising tactile stimuli are produced by a pattern of increased forward-driving effective connectivity within a caudal-to-rostral ascending hierarchy of somatosensory, limbic, and prefrontal areas. Most relevant to embodied predictive coding, we found that within this overall feed-forward flow of causal influences, insula responses to deviants were driven by increases in the strength of both ascending connectivity with prefrontal and cingulate cortex and backwards connectivity with the primary somatosensory area. Importantly, individual differences in the strength of these connectivity modulations predicted participants subjectively rated ease in detecting stimulus changes. This pattern of connections is consistent with the anterior insula acting as core comparator underlying bodily awareness and provides support for core hypothesis of embodied predictive coding.

Previous fMRI studies of oddball responses in the visual, auditory, and tactile modalities report bilateral increases in BOLD activity in the thalamus (TH), primary sensory areas (e.g., S1/V1/A1), anterior insula (AIC), dorso-medial cingulate (MCC), and inferior and middle frontal gyrus (IFG, MFG), all implicated in the present study (See for review: Downar et al. 2002; Garrido et al. 2009). Our mass-univariate results are thus highly consistent with a canonical oddball response in the tactile domain, confirming that unexpected touch is processed in the brain by a coordinated hierarchy of both modality-specific areas (posterior thalamus, S1) and a more cross-modal network of regions likely involved in orienting to salient events (AIC, MCC) and coordinating attention and cognitive control (IFG, MFG).

Deviance responses have been extensively studied using electrophysiological measures which capture the well-characterized mismatch negativity scalp component (Garrido et al. 2009). Studies in the tactile domain have previously demonstrated mismatch responses to sudden changes in stimulus location (Huang et al. 2005), intensity (Chen et al. 2008), and frequency (Kekoni et al. 1997). One previous study using the RSOT modelled the tactile mismatch negativity as encoding Bayesian surprise, a computational marker of prediction error. Interestingly that study found that primary and secondary somatosensory cortices strongly encoded an early (140ms) stimulus-locked rise in Bayesian surprise whereas fronto-insular and cingulate sources showed a later response more associated with the representation of stimulus changes; i.e., salience.

Here we observed strong activation of both S1 and AIC to tactile oddballs, which was mediated by a robust modulation of thalamic afferents to both areas. Both areas were in turn found to directly influence cingulate and prefrontal cortex. A plausible interpretation of both Ostwald's and our own results is that the AIC and MCC jointly monitor the precision or inverse-variance of S1 responses encoding automatic perceptual learning. Precision has been linked computationally to perceptual salience and attention and more specifically to contextual learning via neuromodulatory regulation of post-synaptic cortical gain (Moran et al. 2013). Indeed, the AIC and MCC have been shown to encode to expected precision or volatility (Iglesias et al. 2013; Schwartenbeck et al. 2014) and are generally thought to form part of the 'salience network', facilitating rapid orienting responses to important stimuli. Under predictive coding, salience (i.e. the selection of behaviourally relevant stimuli) can be operationalized as the precision-weighting of prediction errors by post-synaptic modulatory gain (Feldman and Friston 2010; Friston et al. 2012). Consistent with the interpretation that the anterior insula monitors and regulates precision, we found that deviancy signals bypassed S1 to directly modulate the AIC via thalamic afferents, in addition to an indirect route via S1. Thalamic cells are capable of firing in both tonic and 'burst' modes with the latter being important for the processing of salient events (Sherman 2005). Our finding that deviancy directly modulated the AIC, which in turn regulated down-stream S1 responses suggests that the region may monitor the precision of thalamic inputs directly in order to enable fast awareness and responding to critical events (e.g. pain, unexpected touch). This recurrent insular-thalamic-somatosensory loop may be crucial for conscious awareness of tactile changes.

Indeed, recurrent neural activity in the somatosensory hierarchy has previously been shown to be important for conscious somatosensory awareness (Auksztulewicz et al. 2012); in general such cortical-subcortical loops are thought to be critical for conscious awareness (Dehaene et al. 2014). Here we found that strong recurrent connectivity between the AIC and somatosensory cortex supports the processing of tactile oddballs, and that individual differences in the strength of backwards influences from the PFC to AIC predicted the self-rated ease of detecting subtle stimulus changes. These findings together suggest that the AIC coordinates the global cortical processing of surprising bodily stimuli by linking lower-level sensory regions to more attention and salience-related areas in the prefrontal and cingulate cortex during the processing of unexpected tactile changes. As discussed above, an interesting possible interpretation is that the AIC supports the emergence of a global workspace by directly monitoring and modulating the precision of these top-down and bottom-up inputs (Friston and Kiebel 2009; Bastos et al. 2012). Future studies will benefit from directly manipulating tactile precision and deviancy in conjunction with computational modelling to address this question.

Finally, we observed a significant relationship between backwards connection strengths and difficulty ratings, wherein participants whose AIC was most strongly influenced by the PFC also reported the easiest time detecting stimulus changes. As predictive coding postulates that model updates or predictions should be most specifically encoded by top-down connections, this result links top-down effective connectivity or inference to awareness as hypothesized (Bastos et al. 2012). This result thus establishes an essential link between top-down inference, the anterior insula, and embodied awareness, providing criterion validity for our dynamic causal modelling results (Pennington 2003). It is also worth noting that in participants reporting the lowest difficulty discriminating stimulus changes, the presence of strong DLPFC to AIC connectivity effectively completed a cortical-subcortical recurrent loop between somatosensory and prefrontal cortex (compare Figure 4 and 5), in line with a role for the AIC in coordinating a global workspace (Dehaene et al. 2014). However, given that participants merely provided offline ratings which can be subject to a variety of biases, an alternative interpretation may be that participants rated factors unrelated to level of sensory awareness, such as general cognitive effort. We find this interpretation unlikely however; although one previous study did find that DLPFC activations to oddball stimuli are eliminated during a passive oddball task (Clark et al. 2001), the counting task helps to ensure that participants dedicate equivalent attentional effort to deviants and standards, as switches are inherently unpredictable (Garrido et al. 2009). Furthermore, overall switch counts were extremely accurate (mean accuracy = 99%) and participants did not rate either condition as significantly more difficult to detect. These observations make it unlikely that the result is purely confounded by attentional effort, and likely relates to the subjective awareness of intensity differences. Indeed, even with a high level of detection performance and equivalent difficulty ratings across conditions, we observed considerable variability in self-rated difficulty to detect stimulus changes, suggesting that some participants experienced switches more vividly than others. However, to better elucidate the role of attentional control or metacognitive report bias (Fleming and Lau 2014) in this result, it remains an important step for future research to manipulating stimulus probability in the context of a tactile detection task with intensity and perhaps confidence ratings on every trial. Future investigations will help to dissociate the role of executive function and metacognitive awareness in the response to embodied prediction error.

Conclusion

This study demonstrates how tactile awareness recruits a hierarchical mixture of sensory, salience, and attention-related cortical areas to support bodily awareness. Collectively our results illustrate hierarchical processing of surprising tactile stimuli and suggest a critical role for the anterior insula in coordinating global cortical processing and possibly bodily self-awareness. If the AIC does coordinate dynamic interactions between these disparate neural processes via precision modulation, such a mechanism would be crucial for establishing the “global workspace” argued to be necessary for consciousness (Dehaene et al. 2014). Embodied predictive coding may thus provide a framework for understanding how particular predictive codes integrating internal states and external sensory inputs give rise to self-awareness. Understanding how the insula coordinates this complex interaction, and how various sensory channels are integrated and precision-weighted by contextual factors (Feldman and Friston 2010) is likely to yield important future insights into consciousness, emotion, and disruptions thereof. In particular future research may find that hyper- or hypo- connectivity of prefrontal to insula connections plays a critical role in common disorders of bodily awareness, e.g. Ekbom’s syndrome (chronic tactile hallucination) and phantom limb phenomenon.

Acknowledgements

This work was supported by MIND*Lab* Aarhus University (AR, MA, FF), the Danish Agency for Science, Technology and Innovation (MJD), and the Wellcome Trust (MA, GR). The authors thank Karl Friston, Hanneke den Ouden, Benjamin de Haas, Harriet Brown, Chris Frith, and Uta Frith for helpful input on the design, manuscript and analysis.

Tables and Figures Legends

Tables

Table 1

Label	*k*	P_{FWE}	*T*	MNI$_{XYZ}$
R Anterior Insula	403	<0.001	11.16	36 20 1
R Caudate		<0.001	8.74	12 8 4
R Inferior Frontal Gyrus (Area 44)		<0.001	8.15	54 8 19
R Primary Somatosensory Cortex (Area 2)	344	<0.001	10.06	48 -37 46
R Intraparietal Sulcus (hIP1-2)		<0.001	9.65	39 -43 43
R Intraparietal Sulcus (hIP2)		<0.001	9.18	39 -52 52
L Intraparietal Sulcus (hIP1-2)	493	<0.001	9.91	-39 -49 43
L Primary Somatosensory Cortex (Area 2)		<0.001	9.58	-45 -43 49
L Intraparietal Sulcus (hIP1-3)		<0.001	9.13	-33 -55 49
L Pallidum	521	<0.001	9.35	-15 5 4
L Inferior Frontal Gyrus (Area 44)		<0.001	9.10	-48 8 22
L Middle Insula		<0.001	8.08	-42,11,-2
L Anterior Insula		<0.001	8.27	-30,23,-2
L Temporal Gyrus		<0.001	8.49	-48 8 -2
R Middle Frontal Gyrus	160	<0.001	8.45	36 50 22
R Middle Frontal Gyrus		<0.001	7.76	42 44 25
R Middle Frontal Gyrus (DLPFC)		<0.001	7.57	48 32 34
L Anterior Mid-Cingulate	141	<0.001	8.45	-3 20 40
R Anterior Cingulate		0.006	6.62	9 17 25
L Supplementary Motor Area (Area 6)		0.006	6.58	0 14 52
L Thalamus (Th-Prefrontal)	25	<0.001	8.25	-12 -19 10
R Superior Temporal Gyrus	54	<0.001	8.14	48 -22 -5
R Superior Temporal Gyrus		0.002	7.04	48 -31 -5
R Middle Temporal Gyrus		0.012	6.35	57 -37 -5
R Thalamus (Th-Prefrontal)	51	<0.001	8.10	12 -16 10
R Thalamus (Th-Prefrontal)		<0.001	7.49	6 -16 4
L Inferior Parietal Cortex (PF)	35	<0.001	7.86	-57 -43 25
L Middle Temporal Gyrus		<0.001	7.41	-60 -52 16
L Inferior Frontal Gyrus (Area 44)	73	<0.001	7.78	-39 26 28
L Inferior Frontal Gyrus (Area 45)		0.003	6.88	-51 29 28
L Middle Frontal Gyrus		0.008	6.50	-48 38 28
L Inferior Parietal Cortex (PFt, PFop)	33	<0.001	7.59	-57 -22 34
R Middle Frontal Gyrus	39	<0.001	7.36	27 11 58
R Middle Frontal Gyrus		0.010	6.42	39 5 55
L Middle Frontal Gyrus	58	<0.001	7.32	-36 41 22
L Middle Frontal Gyrus		0.002	6.94	-36 50 16
L Prefrontal Gyrus	17	0.002	7.11	-45 2 52
R Superior Temporal Gyrus	32	0.002	7.05	48 -40 13
R Inferior Parietal Cortex (PFcm)		0.012	6.33	57 -40 25
R Inferior Frontal Gyrus (Area 44)	12	0.005	6.67	51 14 40

Figures and Figure Legends

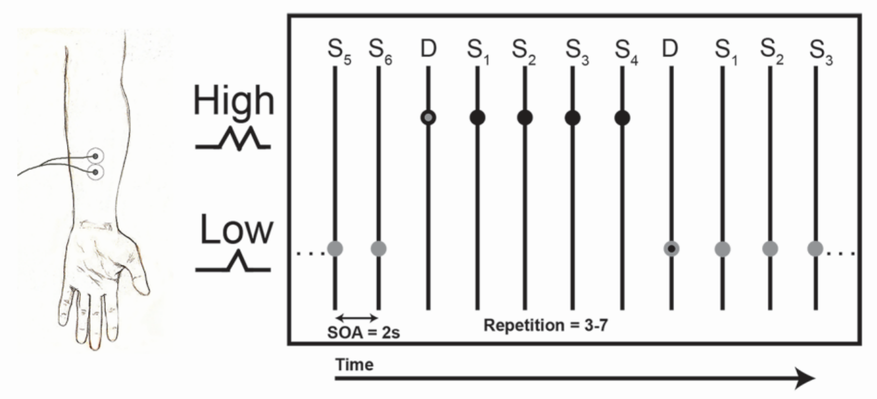

Figure 1. Schematic depicting experimental setup and example stimulus train. Participants received mild somatosensory electrical stimulation (50 µs pulse) at twice sensory threshold on median nerve of the left forearm. Subjective intensity was manipulated by switching between single pulse (bottom-row) and double pulse (top-row) trials. Double pulses were identical to the single pulses, with a 100 ms interstimulus interval. Repetitions varied randomly from 3-7 before switching to the alternate stimulus type, with repetition counts sampled from a random normal distribution. The first stimulus of each train corresponded to a deviant (D), whereas the following repetitions were defined as standards (S1, S2, …, S6). For our fMRI analysis, only the deviant trials and the third standard in each train were modelled (see Methods for more details).

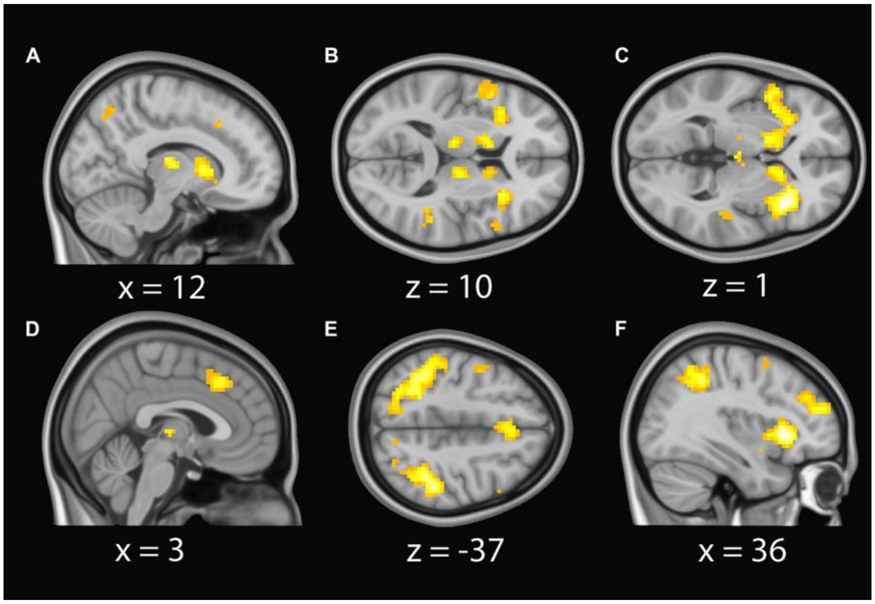

Figure 2. Significant BOLD activations for the deviant > standard contrast. From left to right, images are centred on the peak voxel extracted for each region modelled in the DCM; dorso-posterior thalamus (panels A and B), anterior insula (C), middle cingulate (D), primary somatosensory cortex (E), and the middle frontal gyrus extending into DLPFC (F). Statistical parametric maps, family-wise error corrected for multiple comparisons $P_{FWE} < 0.05$, shown on average of 152 1mm-resolution anatomical scans, normalized to MNI space. Corresponding in-plane MNI coordinate are shown below each image.

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1203v1 | CC-BY 4.0 Open Access | rec: 1 Jul 2015, publ: 1 Jul 2015

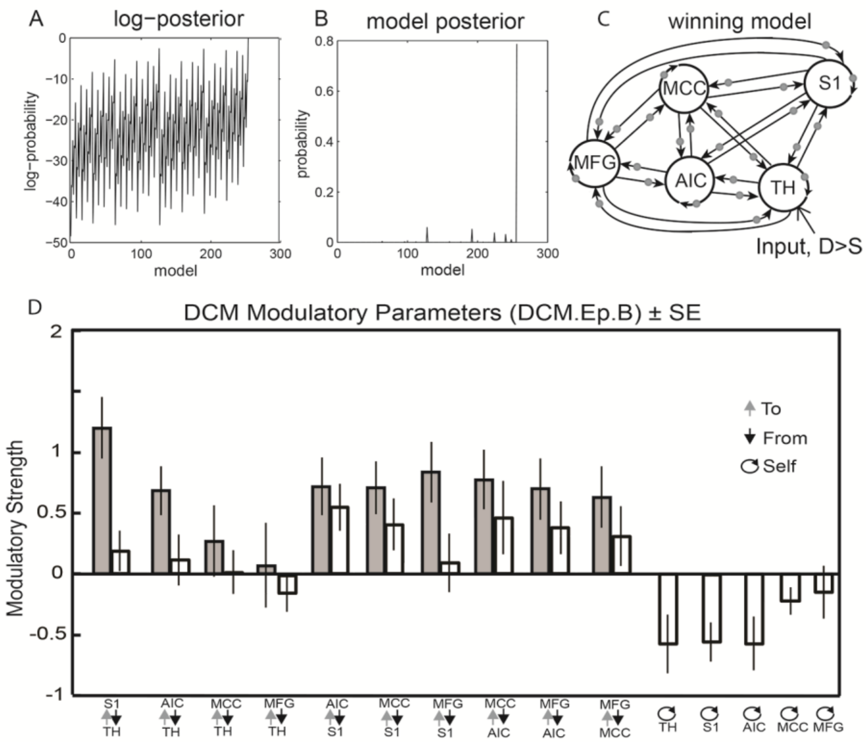

Figure 3. Post-hoc Bayesian model selection (panels A and B), winning model (panel C), and mean coupling parameter plots (panel D). (A) Top left panel depicts the range of log-posterior probability among all models examined. The top middle panel (B) shows the posterior probability for all tested models. Model 255 had the highest probability of 0.79. Model 128 was the next most probable with a posterior probability of 0.06, resulting in a Bayes factor of 13.17 for the full versus reduced model, corresponding to positive evidence that the full model was the best explanation for the measured data within the tested model space. (C) Depiction of the winning full model (Model 255, far right peak in Figure 3B), gray circles indicate modulation by the Deviant > Standard contrast. (D) Bar plot depicting mean posterior parameter estimates for all modulatory (DCM.Ep.B) parameters across subjects, indicating the strength in Hertz with which each connection was modulated by deviant > standard stimuli. Error bars depict standard error. Modulations of inhibitory self-connections are shown at the right hand side of the graph.

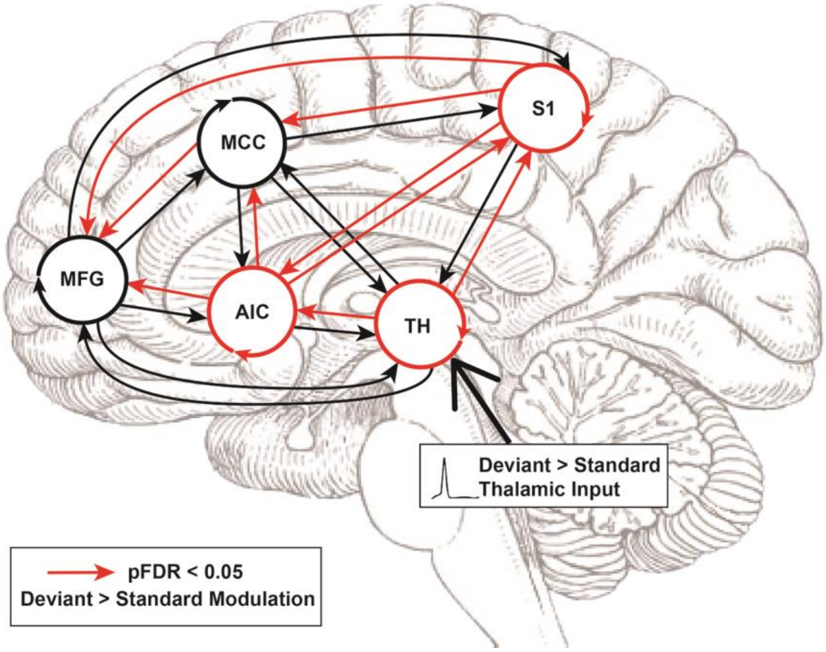

Figure 4. Full model and results of one-sample t-tests over estimated modulatory parameters. Red arrows depict results of one-sample t-tests over all 25 modulation parameters (inhibitory self-connections indicated by circular arrow around each region label). A general caudal to rostral flow increased effectivity connectivity in response to tactile deviants can be observed from thalamus (TH), and primary somatosensory cortex (S1), to anterior insula (AIC) and mid-cingulate (MCC), before reaching prefrontal cortex (middle frontal gyrus, MFG). In contrast to this feed-forward flow of modulatory influences, the AIC shows significant increases in both 'forwards' connections to cingulate and prefrontal cortex and 'backwards' connections with S1, indicative of error comparison. Interestingly, TH, AIC, and S1 self-connections are strongly disinhibited by tactile deviants. All p-values false discovery rate corrected for multiple comparisons, $P_{FDR} < 0.05$.

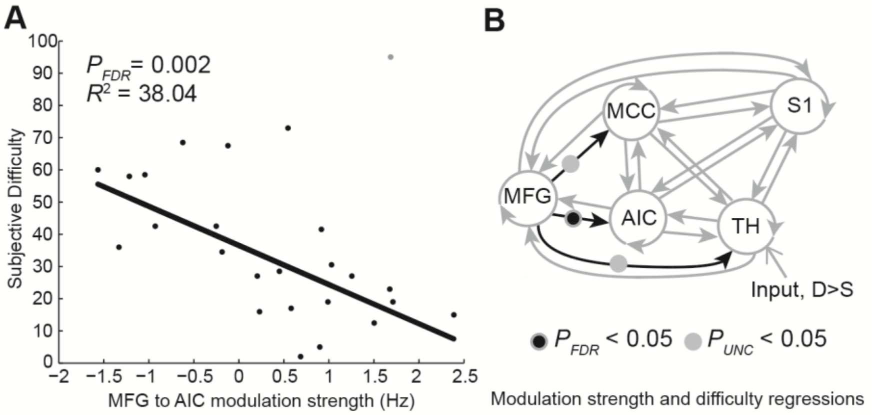

Figure 5. Robust Regression analysis with self-reported difficulty for detecting stimulus changes predicting the strength of deviance-driven modulation of effective connectivity. Left panel A, participants with enhanced modulation of the backwards MFG to AIC connection by surprising touch stimuli reported easier discrimination of stimulus changes. Data points depict individual participants; points shaded gray indicate those receiving down weights > 2 SD from the mean weighting (leverage points). These results suggest that top-down prefrontal to AIC connectivity underlies awareness of unexpected tactile changes, as predicted by embodied predictive coding. Right panel B depicts results of robust regressions (Tukey's biweight) over 20 intrinsic connection modulation parameters each predicting subjective difficulty, P_{FDR} threshold < 0.05.

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1203v1 | CC-BY 4.0 Open Access | rec: 1 Jul 2015, publ: 1 Jul 2015

References

Andersson JLR, Hutton C, Ashburner J, Turner R, Friston K. 2001. Modeling Geometric Deformations in EPI Time Series. Neuroimage. 13:903–919.

Apps M a J, Tsakiris M. 2013. The free-energy self: A predictive coding account of self-recognition. Neurosci Biobehav Rev.

Ashburner J, Friston KJ. 2005. Unified segmentation. Neuroimage. 26:839–851.

Auksztulewicz R, Spitzer B, Blankenburg F. 2012. Recurrent neural processing and somatosensory awareness. J Neurosci. 32:799–805.

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

Brown H, Adams R, Parees I, Edwards M, Friston K. 2013. Active inference, sensory attenuation and illusions. Cogn Process. 14:411–427.

Büchel C, Geuter S, Sprenger C, Eippert F. 2014. Placebo Analgesia: A Predictive Coding Perspective. Neuron. 81:1223–1239.

Cerliani L, Thomas RM, Jbabdi S, Siero JCW, Nanetti L, Crippa A, Gazzola V, D’Arceuil H, Keysers C. 2012. Probabilistic tractography recovers a rostrocaudal trajectory of connectivity variability in the human insular cortex. Hum Brain Mapp. 33:2005–2034.

Chang LJ, Yarkoni T, Khaw MW, Sanfey AG. 2012. Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cereb Cortex. bhs065.

Chen TL, Babiloni C, Ferretti A, Perrucci MG, Romani GL, Rossini PM, Tartaro A, Del Gratta C. 2008. Human secondary somatosensory cortex is involved in the processing of somatosensory rare stimuli: an fMRI study. Neuroimage. 40:1765–1771.

Clark VP, Fannon S, Lai S, Benson R. 2001. Paradigm-dependent modulation of event-related fMRI activity evoked by the oddball task. Hum Brain Mapp. 14:116–127.

Craig A. 2003. Interoception: the sense of the physiological condition of the body. Curr Opin Neurobiol. 13:500–505.

Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ. 2004. Neural systems supporting interoceptive awareness. Nat Neurosci. 7:189–195.

Dehaene S, Charles L, King J-R, Marti S. 2014. Toward a computational theory of conscious processing. Curr Opin Neurobiol. 25C:76–84.

Dietz MJ, Friston KJ, Mattingley JB, Roepstorff A, Garrido MI. 2014. Effective connectivity reveals right-hemisphere dominance in audiospatial perception: implications for models of spatial neglect. J Neurosci. 34 :5003–5011.

Downar J, Crawley AP, Mikulis DJ, Davis KD. 2002. A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. J Neurophysiol. 87:615–620.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data, NeuroImage.

Feldman H, Friston K. 2010. Attention, uncertainty and free-energy . Front Hum Neurosci .

Fleming SM, Lau HC. 2014. How to measure metacognition. Front Hum Neurosci. 8:1–9.

Friston K, Adams R a., Perrinet L, Breakspear M. 2012. Perceptions as hypotheses: Saccades as experiments. Front Psychol. 3:1–20.

Friston K, Kiebel S. 2009. Predictive coding under the free-energy principle. Philos Trans R Soc Lond B Biol Sci. 364:1211–1221.

Friston K, Penny W. 2011. Post hoc Bayesian model selection. Neuroimage. 56:2089–2099.

Friston KJ, Li B, Daunizeau J, Stephan KE. 2011. Network discovery with DCM. Neuroimage. 56:1202–1221.

Gallagher S. 2005. How the body shapes the mind. Cambridge Univ Press.

Garrido MI, Friston KJ, Kiebel SJ, Stephan KE, Baldeweg T, Kilner JM. 2008. The functional anatomy of the MMN: a DCM study of the roving paradigm. Neuroimage. 42:936–944.

Garrido MI, Kilner JM, Stephan KE, Friston KJ. 2009. The mismatch negativity: a review of underlying mechanisms. Clin Neurophysiol. 120:453–463.

Gu X, Hof PR, Friston KJ, Fan J. 2013. Anterior insular cortex and emotional awareness. J Comp Neurol. 3388:3371–3388.

Huang M-X, Lee RR, Miller G a, Thoma RJ, Hanlon FM, Paulson KM, Martin K, Harrington DL, Weisend MP, Edgar JC, Canive JM. 2005. A parietal-frontal network studied by somatosensory oddball MEG responses, and its cross-modal consistency. Neuroimage. 28:99–114.

Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R. 2002. Image Distortion Correction in fMRI: A Quantitative Evaluation. Neuroimage. 16:217–240.

Hutton C, Deichmann R, Turner R, Andersson JLR. 2004. Combined correction for geometric distortion and its interaction with head motion in fMRI. In: Proceedings of ISMRM. p. 1084.

Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, Stephan KE. 2013. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. Neuron. 80:519–530.

Kekoni J, Hämäläinen H, Saarinen M, Gröhn J, Reinikainen K, Lehtokoski A, Näätänen R. 1997. Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. Biol Psychol. 46:125–142.

Keltner JR, Furst A, Fan C, Redfern R, Inglis B, Fields HL. 2006. Isolating the modulatory effect of expectation on pain transmission: a functional magnetic resonance imaging study. J Neurosci. 26:4437–4443.

Lieder F, Stephan KE, Daunizeau J, Garrido MI, Friston KJ. 2013. A Neurocomputational Model of the Mismatch Negativity. PLoS Comput Biol. 9:e1003288.

Lohmann G, Erfurth K, Müller K, Turner R. 2012. Critical comments on dynamic causal modelling. Neuroimage.

Lovero KL, Simmons AN, Aron JL, Paulus MP. 2009. Anterior insular cortex anticipates impending stimulus significance. Neuroimage. 45:976–983.

Markov NT, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Toroczkai Z, Kennedy H. 2013. Cortical High-Density Counterstream Architectures. Sci . 342 .

Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ. 2013. Free energy, precision and learning: the role of cholinergic neuromodulation. J Neurosci. 33:8227–8236.

Nichols T, Hayasaka S. 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res. 12:419–446.

O’Doherty JP, Deichmann R, Critchley HD, Dolan RJ. 2002. Neural responses during anticipation of a primary taste reward. Neuron. 33:815–826.

Ostwald D, Spitzer B, Guggenmos M, Schmidt TT, Kiebel SJ, Blankenburg F. 2012. Evidence for neural encoding of Bayesian surprise in human somatosensation. Neuroimage. 62:177–188.

Peirce JW. 2007. PsychoPy—Psychophysics software in Python. J Neurosci Methods. 162:8–13.

Pennington DC. 2003. Essential personality. Oxford University Press.

Penny WD, Stephan KE, Mechelli A, Friston KJ. 2004. Comparing dynamic causal models. Neuroimage. 22:1157–1172.

Poldrack R a. 2012. The future of fMRI in cognitive neuroscience. Neuroimage. 62:1216–1220.

Schwartenbeck P, FitzGerald THB, Mathys C, Dolan R, Friston K. 2014. The Dopaminergic Midbrain Encodes the Expected Certainty about Desired Outcomes. Cereb Cortex. 1–12.

Serino A, Haggard P. 2010. Touch and the body. Neurosci Biobehav Rev. 34:224–236.

Seth AK. 2013. Interoceptive inference, emotion, and the embodied self. Trends Cogn Sci. 17:1–21.

Seth AK, Suzuki K, Critchley HD. 2011. An interoceptive predictive coding model of conscious presence. Front Psychol. 2:395.

Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R. 2005. Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. Nat Neurosci. 8:1234–1240.

Sherman SM. 2005. Thalamic relays and cortical functioning. Prog Brain Res. 149:107–126.

Ullsperger M, Harsay H a, Wessel JR, Ridderinkhof KR. 2010. Conscious perception of errors and its relation to the anterior insula. Brain Struct Funct. 214:629–643.