

Confidence of emotion expression recognition recruits brain regions outside the face perception network

Indrit Bègue,^{1,2,3} Maarten Vaessen,¹ Jeremy Hofmeister,¹ Marice Pereira,¹ Sophie Schwartz,^{1,2} and Patrik Vuilleumier^{1,2}

¹Department of Neuroscience, Faculty of Medicine, University of Geneva, CH-1211, Switzerland, ²Swiss Center for Affective Studies, University of Geneva, CH-1202, Switzerland, and ³Department of Mental Health and Psychiatry, University Hospitals of Geneva, CH-1226, Switzerland

Correspondence should be addressed to Indrit Bègue, Department of Neuroscience, Faculty of Medicine, University of Geneva, 1 Rue Michel Servet 1211 Geneva 4, Switzerland. E-mail: indrit.sinanaj@unige.ch.

Abstract

Metacognitive beliefs about emotions expressed by others are crucial to social life, yet very little studied. To what extent does our confidence in emotion expression recognition depend on perceptual or other non-perceptual information? We obtained behavioral and magnetic resonance imaging measures while participants judged either the emotion in ambiguous faces or the size of two lines flanking these faces, and then rated their confidence on decision accuracy. Distinct behavioral and neural mechanisms were identified for confidence and perceptual decision in both tasks. Participants overestimated their emotion recognition (ER) accuracy, unlike visual size judgments. Whereas expression discrimination recruited several areas in the face-processing network, confidence for ER uniquely engaged the bilateral retrosplenial/posterior cingulate complex and left parahippocampal gyrus. Further, structural white matter connectivity of the former region predicted metacognitive sensitivity. These results highlight a key role for brain mechanisms integrating perception with contextual mnemonic information in the service of confidence during ER.

Key words: metacognition; confidence; emotion recognition; fMRI

Introduction

I know how you feel; or do I really? Confident but wrong beliefs about the emotion felt or expressed by others may be damaging to social relationships. Albeit highly complex and temporally unfolded in some situations, social encounters are often guided by rapid appraisals of simple sensory cues—think of a stranger with whom you fleetingly cross eyes in a night train. Humans have evolved sophisticated mechanisms for quick and efficient recognition of emotion expression from faces (Haxby *et al.*, 2000). However, even though it is vital to swiftly recognize facial expressions, beliefs about how we trust what the face is conveying ('I am sure he is angry at me') are also important to adjust behavior, contributing to either provoke or avoid fatal misunderstandings.

Abilities by which we form beliefs on the working of our own mind are called metacognitive, and the process, metacognition (Flavell, 1976; Fleming and Dolan, 2012). Metacognitive abilities reflect the credibility of these beliefs, i.e. how faithfully our self-judgment relates to actual performance in a task and can be measured by retrospective confidence ratings. Research in neuroscience has only begun to unravel the neural mechanisms of these self-monitoring processes and how they relate to actual discrimination performance. However, previous research has generally focused on elementary perceptual skills and memory (Grimaldi *et al.*, 2015). Virtually nothing is known about self-monitoring processes for recognition of emotions in others, despite intense research on social cognition and face perception.

Received: 20 July 2018; Revised: 1 November 2018; Accepted: 21 November 2018

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work properly cited. For commercial re-use, please contact journals.permissions@oup.com

Moreover, it remains debated whether perceptual decisions and confidence share neural and computational substrates (Kepecs et al., 2008; Kiani and Shadlen, 2009), or whether metacognitive judgments rely on specific brain circuitry mainly located in prefrontal regions (for reviews see Fleming and Dolan, 2012; Grimaldi et al., 2015; Yeung and Summerfield, 2012). According to an influential account, confidence judgments are directly derived from signals mediating perceptual decisions (Kepecs et al., 2008; Kiani and Shadlen, 2009). By contrast, Shadlen and Shohamy (2016) recently advocated a new hypothesis stipulating that more evaluative decisions (i.e. choosing between two food items), unlike elementary perceptual decisions (i.e. evaluating whether dots move towards the left or right), may involve retrieval of information from memory (past experiences) in addition to purely perceptual information. However, there is still scarce evidence showing that confidence requires the integration of sensory signals beyond those used for the perceptual decision process itself. Accordingly, emotion recognition (ER) may differ from simple perceptual tasks since in the former case subjects may draw evidence based on multiple sensory cues as well as context and past experiences stored in memory.

Here we designed a novel paradigm to directly assess the behavioral parameters and brain systems underlying metacognition of facial ER and of visual size perception (S) judgments. Specifically, we asked two key questions that remain unresolved (Fleming and Dolan, 2012): (i) Does self-confidence in ER ability recruit distinct or similar metacognitive mechanisms as compared to those reported for other visual discrimination tasks? (ii) Do ER-related and perception-related metacognitive processes rely on brain functions that support recognition abilities (e.g. extrastriate visual areas), or do they involve higher-level systems that integrate visual information with other cognitive/affective functions (e.g. medial prefrontal regions)? Additionally, as an auxiliary question, we were interested to explore whether particular interindividual differences in empathy (capacity to read other people mental states), alexithymia (capacity to identify and describe emotions), or narcissism (propensity to overconfidence) would relate to the variability of metacognitive performance for ER.

To answer these questions, we asked participants ($n = 34$) to either judge the emotion displayed in a morph of happy and angry faces, or report the thicker of two lines at the top and bottom of these faces (Figure 1). On each trial, participants rated their confidence in these judgments. We could thus directly compare metacognition for ER with another visual perceptual discrimination task. Crucially, we introduced a jittered non-informative feedback screen between the decision and the confidence rating, which allowed us to compare neural activity associated with these two events. Finally, we examined how metacognitive performance in each task related to individual variability in standard personality questionnaires that have been related to trait differences in ER and recognition confidence (Kelly and Metcalfe, 2011).

Materials and methods

Participants

Thirty-four healthy volunteers (15 females, mean age = 23.8 years old, range = 19–32) were recruited by online announcements and written advertisement. They were screened for psychiatric or neurological disorders and had normal or corrected-to-normal vision. All participants were included in the func-

tional magnetic resonance imaging (fMRI) analyses. The study was approved by the Ethics Committee of the University of Geneva and University Hospitals of Geneva. All participants read and signed the informed consent form.

ER and visual perception tasks

The experimental task is illustrated in Figure 1. Each trial consisted of a single face presented with bars on the top and bottom (duration = 200 ms; visual angle 4.9° horizontally \times 8.9° vertically) (Figure 1). Participants had to either categorize the face expression (angry/happy) or indicate the thicker of the two lines (top/bottom) in a two-alternative forced-choice fashion (response time window = 3000 ms; see Supplementary Material for details). We used four face identities, in which angry emotional expressions (image 1) were morphed across 80 steps towards happy emotional expression (image 80). The 40th image obtained held equal image information from the angry and happy emotional expressions. The two tasks were given in blocks (eight trials), preceded by an instruction screen (i.e. 'Judge Emotion' or 'Judge Thickness') indicating which task to perform. There were a total of 256 trials with 128 trials per each condition (ER, S). Faces used for the ER task in the n block (with expressions determined by the ER staircase) were then used for the S task in the $n + 1$ block (and vice versa for bars for the $n + 2$ block, etc.), allowing us to present similar stimuli during perceptual and ER judgments. The order of the blocks was counterbalanced between subjects. Task performance was kept fixed in both tasks via an adaptive transformed and weighted one-up-two-down staircase procedure per task (Fleming et al., 2010). This aimed to ensure globally similar performance between conditions and across individuals, in order to rule out any confounding influence of relative task difficulty on confidence judgments. Finally, participants had to rate their confidence level on the accuracy of their previous discrimination (response time window = 4000 ms).

Participants made two judgments on each trial (Figure 1) after being prompted by two successive response screens: first, stimulus discrimination (ER task: happy vs angry; perception task: top vs bottom) and second, subjective confidence of their response on this trial (from 1 = totally unsure, to 6 = totally certain, always starting at the middle of the scale corresponding to mean confidence level = 3.5). Two buttons were used to move the cursor leftward or rightward, and a third to confirm the response. Participants received no feedback on their response accuracy during the whole duration of the experiment.

Behavior analysis

Metacognition quantification

Metacognitive performance is characterized by calibration (i.e. how well confidence tracks accuracy) (Adams and Adams, 1961) and resolution (i.e. how well subjects discriminate between their correct and incorrect decisions). Confidence ratings may distinguish between correct and incorrect responses (e.g. have good resolution) yet be poorly calibrated if mean probability of being correct outstrips actual performance (Fleming and Lau, 2014). Overconfidence occurs when the expressed confidence level exceeds accuracy, and conversely, underconfidence occurs when expressed confidence is lower than accuracy. Here, we linearly transformed the raw confidence ratings (that included seven levels of confidence: 1, 2, 3, 3.5, 4, 5 and 6) into a scale from 0–1 (0 corresponds to the lowest and 1 to the highest confidence),

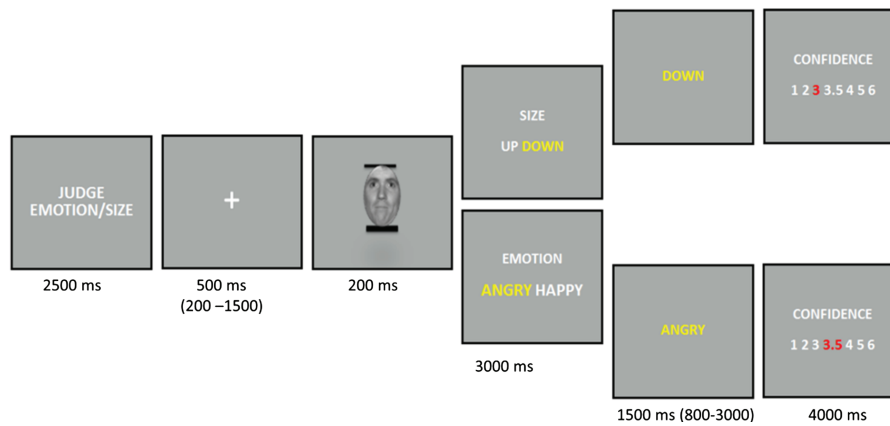


Fig. 1. Overview of experimental events. A written cue ('Judge Emotion' or 'Judge Thickness') appeared at the beginning of each block to indicate the task to be performed during the next trials. Each trial began with a fixation cross (mean 500 ms, jittered between 200–1500 ms), followed by a face (200 ms) with two bars. The face was obtained by morphing its respective happy and angry expressions in 80 steps (the intermediate 40th face image is displayed here). Participants had to decide whether the face was angry or happy during ER blocks, or whether the thicker bar was on the top or bottom during the S blocks. After each answer, they rated how confident they felt about their judgment on this particular trial from 1 ('not certain at all') to 6 ('very certain'). No feedback on performance accuracy was provided. However, to separate events of interest ('discrimination' and 'confidence') for our fMRI analysis, the participant's response was displayed in the interval (mean duration 1500 ms, jittered between 800–3000 ms), followed by the confidence rating screen. In order to approximately keep the same total duration for all trials, the sum of the remaining time relative to the maximal response durations from the discrimination and confidence screens was added to the next fixation cross screen.

then calculated an over/underconfidence (O/UC) score as the difference between the mean confidence level and the mean accuracy for each individual (Fischhoff et al., 1977):

$$\text{over/underconfidence score} = x - c,$$

where x is the mean confidence score and c is the mean proportion correct (for details, see discussion in Pulford, 1996).

Resolution of metacognitive ability can be quantified by measuring the area under the receiver operating characteristic curve (A_{ROC}) type 2 using non-parametric methods (Fleming and Lau, 2014), see [Supplementary Material](#) for a full description. ROC curves type 2 can evaluate how well the subject's confidence rating can discriminate between correct and incorrect decisions (Galvin et al., 2003). However, one limitation of A_{ROC} is that it can be influenced by task performance. Here we controlled performance in both tasks via adaptive transformed and weighted one-up-two-down staircase procedures (one per task) (Fleming et al., 2010), although our staircases failed to completely equate performance between tasks (see Results). Metacognitive sensitivity, as thus quantified by the A_{ROC} , has meaningful interpretative values. When A_{ROC} is equal to zero, this indicates that participant's confidence is unable to discriminate between correct and incorrect responses; therefore, the observer displays no metacognitive sensitivity. Higher A_{ROC} values indicate better metacognitive sensitivity (Fleming and Lau, 2014). We also calculated the meta- d' (Maniscalco and Lau, 2012), which is a measure of metacognitive sensitivity, using analysis' scripts used in previous work and publicly available (<http://www.columbia.edu/~x007E/bsm2105/type2sdt/>). Further, we calculated an M_{ratio} (meta- d'/d') which examines the efficacy of metacognitive mechanisms, i.e. how well task-related information is utilized for metacognition (Maniscalco and Lau, 2012) unconfounded by type I performance.

Questionnaires

To address interindividual trait differences in metacognition and ER recognition, we collected a series of standardized

questionnaires. Specifically, we examined whether the capacity to accurately infer other people's emotions (empathy), to describe one's own emotions (alexithymia), and to have a general propensity to being overconfident (narcissism) could affect metacognition (Kelly and Metcalfe, 2011). Empathy was measured using the Balanced Emotional Empathy Scale (BEES) (Mehrabian, 1996), the Empathy Quotient (EQ) of Baron-Cohen (Baron-Cohen and Wheelwright, 2004) and the Interpersonal Reactivity Index (IRI) (Davis, 1980). Additionally, we administered the 20-item Toronto Alexithymia Scale (TAS-20) (Bagby et al., 1994) and the O'Brien Multiphasic Narcissism Inventory (OMNI) (O'Brien, 1987) for alexithymia and narcissism, respectively. Lastly, we collected state anxiety scores using the State-Trait Anxiety Inventory (Spielberger, 1983).

Image acquisition and analysis

MRI data were acquired using a 3T whole-body MRI scanner (Trio TIM, Siemens, Germany) with the product 32-channel head coil. For each participant, functional images were acquired with an echo-planar imaging (EPI) sequence (TR/TE/flip angle = 2100 ms/30 ms/80°, slice thickness = 3.2 mm, voxel size = 3.2 × 3.2 × 3.2 mm). For anatomic reference, a T1-weighted 3D spoiled fast gradient echo pulse sequence was acquired with the following parameters: TR/TI/TE/flip angle = 1900 ms/900 ms/2.27 ms/98°, FOV = 256 mm, voxel size = 0.9 mm isotropic). Diffusion weighted images (DWI) were also obtained in the same session: pixel size = 1.5 × 1.5 mm², slice thickness 1.5 mm and a B-value of 1500 s/mm². Diffusion weighted volumes were acquired at 65 unique gradient directions interspersed with six non-diffusion weighted volumes (Jones, 2004). An EPI sequence was used with TE 109 ms, TR 5163 ms and a multi-band acceleration factor of 3 (Setsompop et al., 2012). The DWI acquisition time was 5.5 min.

Statistical analysis was performed using the SPM8 software package (<http://www.fil.ion.ucl.ac.uk/spm/>). Scans from all four sessions were realigned (Friston et al., 1995), then co-registered to the anatomical image, normalized to the EPI template (resampled at 3 × 3 × 3 mm voxel size) and spatially smoothed (8 mm

full width at half maximum Gaussian kernel). Statistical analysis was performed on a voxel-wise fashion across the whole brain. Individual events were modeled by a standard hemodynamic response function. We modeled four main event types (Model 1): the onset of the visual stimulus (face and bars) to be discriminated according to the relevant task (DISC) and the onset of the confidence screen (CONF) where subjects evaluated their performance, for both the ER (E) and S (S) tasks (i.e. four conditions: DISC S, DISC E, CONF S and CONF E). We included two parametric modulators for the discrimination period [confidence level: 1–7; normalized stimulus intensity: 0 minimal–1 maximal stimulus intensity], and two modulators for the confidence period [confidence level: 1–7; and number of key presses to account for the fact that highest/lowest confidence ratings required more key presses compared to midscale confidence levels]. The task instruction screen (i.e. 'Judge Emotion' or 'Judge Thickness') was presented only at the beginning of each task block (see above) and not modelled explicitly. To account for residual movement-related variance, we also included six differential movement parameters derived from pre-processing as regressors of no interest. We also created a second model (Model 2) where reaction times were added as a parametric modulator of the confidence rating regressors of Model 1 (CONF), which yielded results similar to Model 1.

We performed direct contrasts between DISC E > DISC S and between DISC S > DISC E to identify brain regions differentially involved in ER and visual size perception, respectively. Brain regions linked with confidence judgments for ER were determined by the contrast CONF E > CONF S, and vice versa for confidence for S (CONF S > CONF E). Note that during DISC and CONF events, participants performed two successive judgments concerning the same stimulus, differing in that CONF (type 2 judgment) but not DISC (type 1 judgment) required explicit metacognitive representations.

To distinguish between these two events at the neural event, we introduced a pseudo-randomized jittered-duration blank interval between the discrimination and confidence response screens (mean = 1500 ms, range = 800–3000 ms). This manipulation resulted in a mean temporal distance between the CONF and DISC screens of 6 s (s.d. = 0.90 s) for the ER task and 5.91 s (s.d. = 0.84 s) for the S task (See [Figure S1](#), in the [Supplementary Material](#) for mean values across subjects of the trial-to-trial duration of the jitter in seconds between DISC and CONF). Although jittering may not fully separate blood-oxygen-level dependent (BOLD) signal changes associated with each process, this duration is similar to stimulus-onset asynchrony (SOA) used in standard event-related studies, allowing us to reliably uncorrelate the corresponding regressors and thus highlight brain areas differentially involved in discrimination and confidence judgments for each task domain (Josephs and Henson, 1999). Crucially for our experiment, we computed the interaction contrast of task (ER vs S) x judgment (CONF vs DISC) to reveal task-specific activations related specifically to metacognitive confidence (but not discrimination) in ER (CONF E > DISC E) > (CONF S > DISC S) and S (CONF S > DISC S) > (CONF E > DISC E). For common activations related to the metacognition of ER and S we computed the conjunction null hypothesis (CONF E > DISC E) AND (CONF S > DISC S) testing for a logical AND. In *post hoc* analyses, we examined whether activity in brain regions directly involved in perceptual discrimination (DISC E > DISC S, DISC S > DISC E, respectively) or confidence judgment (CONF E > CONF S, CONF E > CONF S, respectively), for either task, correlated with metacognitive indices (A_{ROC} and M_{ratio} , statistical threshold of $P < 0.05$) by using a region of interest

(ROI) approach with a sphere of 5 mm over peak voxels defined by the abovementioned contrasts. For completeness and in line with other studies in the field (Fleming et al., 2012; Morales et al., 2018), metacognitive indices (A_{ROC} and M_{ratio}) were also used as covariates in second-level analyses during the same contrasts (main effects of confidence, CONF E > CONF S and vice versa; and interaction effect, (CONF E > DISC E) > (CONF S > DISC S) and vice versa), allowing us to identify any parametric effects at the whole-brain level (with statistical threshold of $P < 0.001$ uncorrected at the voxel peak).

All contrasts arising from the single-subject first-level analyses above were fed into one-sample t-tests to determine significant group effects. Significant brain activations maps were thresholded at $P < 0.001$ (uncorrected) with significant clusters corrected at the cluster level [$P < 0.05$; family-wise error (FWE)]. The unthresholded statistical maps from the second level analyses can now be found at <https://neurovault.org/collections/UUVU-RAGIG/>.

With respect to the structural diffusion tensor images (DTI) analysis, spatial processing of the DTI with motion and eddy current correction, brain extraction and normalization, as well as calculation of tensor and fractional anisotropy (FA) values were performed with the FMRIB Software Library (FSL) routines (Jenkinson et al., 2012). First, individual FA maps were normalized to the FSL MNI space using the standard FA map template. Second, a whole-brain general linear model (GLM) with AROC values as regressor was used to test for any association between FA values and AROC scores for ER and perception. Maps of z-values obtained from this regression analysis were then analyzed for clusters of significant associations using the FSL cluster method ($P < 0.05$).

Results

Behavior

While the difficulty of each was adjusted individually using a dynamic online staircase procedure, visual S discrimination was slightly but significantly better overall compared to ER (mean correct = 82%, s.d. = 3 vs 79%, s.d. = 2, respectively, $T = 4.95$; mean $d_{\text{prime}} = 1.74$, s.d. = 0.77 vs 1.56, s.d. = 0.17, respectively, $T = 5.48$, $P < 0.001$, [Figure 2A](#)). Nevertheless, participants were much more confident of their ER judgments (mean = 5.88, s.d. = 0.49 on a scale of seven levels) as compared to visual size (S) judgments (mean = 4.95, s.d. = 0.78, $t = 10.11$, $P < 0.001$). Importantly, this difference in confidence remained true even when controlling for differences in task performance ($F = 19.47$, $P < 0.001$) or d_{prime} ($F = 16.9$, $P < 0.001$) with a multivariate analysis of covariance (MANCOVA) so it is correct ([Figure 2C](#)). Thus, despite the fact that ER tended to be harder than S judgments, participants overestimated their ability in the former relative to the latter with no difference in metacognitive bias (B_{ROC} ; i.e. overall propensity to provide higher or lower confidence ratings) between these two judgments (ER: mean = -0.23 , s.d. = 0.4, S : mean = -0.14 , s.d. = 0.61, $T = 0.65$, n.s.; [Figure 2B](#)).

As expected, performance was better on trials with more distinctive stimuli: there was a positive relationship between stimulus intensity level (expression morphing or line thickness difference) and accuracy for both ER (Spearman -0.57 , $P = 0.007$) and S decisions ($r = 0.67$, $P = 0.002$; [Figure 2D and F](#)). We also found a positive relationship between stimulus intensity and confidence level for size discrimination (Spearman $r = 0.47$, $P = 0.04$) ([Figure 2G](#)). However, this correlation was not significant for ER ($r = 0.21$, n.s.) because confidence tended to be

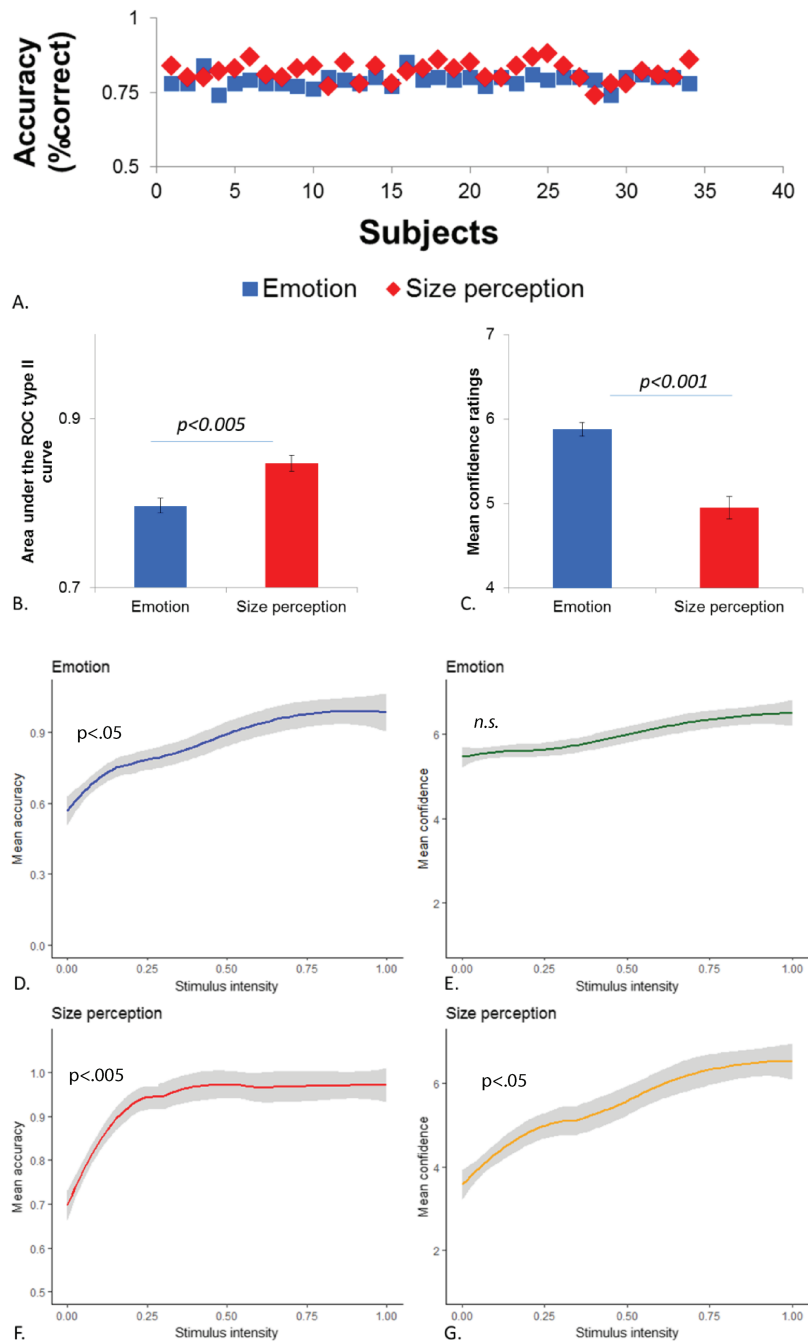


Fig. 2. Overview of behavioral results. Accuracy scores (percent correct) for all subjects (A), mean area under the ROC Type II curve (B) and mean confidence ratings (C) for ER (blue) and S (red) tasks. Errors bars represent standard error. Relationship of the stimulus intensity level with (D, F) mean accuracy and (E, G) mean confidence ratings, for each task (ER, left panels; S, right panels) and accompanying 95% confidence bands in the shaded area. Accuracy progressively increased with increasing stimulus intensity levels for both tasks. Mean confidence tracked well stimulus intensity levels for the S task, but remained relatively high and flat throughout the different stimulus levels for ER.

high even at low stimulus intensity levels (Figure 2E). These results point to distinct processes underlying the computation of confidence for S and ER performance, with the former being proportional to stimulus intensity whereas the latter may at least partly rely on some other information (e.g. previous experience, prior beliefs), possibly arising outside purely perceptual processes.

We then computed O/UC scores for both tasks. For the S task, the average O/UC score did not differ from zero, implying good calibration of confidence judgments; but remarkably, for ER, the

O/UC score was significantly above zero (mean = 0.09; $T = 9.26$, $P < 0.001$), implying poor calibration and overconfidence (i.e. inflated subjective trust in performance).

Individual metacognitive sensitivity was also quantified by calculating the area under the ROC curve (A_{ROC}) for each task (see Supplementary Material). These data indicated lower metacognitive sensitivity for ER compared to S judgments (A_{ROC} for ER: mean = 0.80, s.d. = 0.04, A_{ROC} for S: mean = 0.84, s.d. = 0.05, $T = 3.71$, $P = 0.001$; Figure 2B). A regression analysis with task (ER vs S) as independent factor and A_{ROC} values as dependent

variable showed a highly significant main effect of task (ER vs S; $F = 14.33$, $P < 0.001$; $F = 4.20$, $P = 0.002$, corrected for performance as indexed by percent correct for each task). A repeated measures t-test also revealed significant differences for the meta- d' (i.e. the theoretical d' that would explain observed confidence ratings; ER mean = 1.58, s.d. = 0.49; S mean = 2.05, s.d. = 0.47, $T = 4.2$, $P < 0.001$). However, Mratio (a metacognitive index weighted by differences in first-order discrimination performance, $metad' - d'$) showed no significant differences between ER (mean = 1.04, s.d. = 0.35) and S (mean = 1.09, s.d. = 0.22) judgments ($T = 0.70$, n.s.). Importantly, metacognitive abilities in one task did not predict metacognitive abilities in the other (Spearman r between A_{ROC} for ER and S, $r = 0.06$, n.s., Spearman r between Mratio for ER and S, $r = 0.07$, n.s.), i.e. subjects with high metacognition for ER were not necessarily those with high metacognition for visual size perception.

Taken together, these behavioral results indicate that while confidence in size discrimination tracks well actual performance and stimulus intensities, participants appear overconfident of their ER ability, particularly for the most ambiguous expressions. This suggests worse estimation of subjective confidence for ER. However, this might at least partially be related to lower performance at the discrimination level, as the weighted metacognitive indices (Mratio) were similar in both tasks. Alternatively, or additionally, this pattern might reflect reliance on different sources of information to judge confidence of performance during emotion and visual recognition tasks.

Individual differences associated with metacognition

Finally, we examined whether ER performance and metacognitive sensitivity correlated with individual differences in empathy skills as assessed by standard questionnaires [EQ, the Empathic Concern subscale of IRI (IRI-EC) and BEES]. We found a consistent positive relation of empathy with A_{ROC} values in the ER task (Spearman $r = 0.38$, $P = 0.03$ and $r = 0.36$, $P = 0.04$, for IRI-EC and BEES, respectively) but not in the S task (Spearman $r = -0.13$ and $r = -0.16$, all P -values n.s.). Correlations with EQ were not significant (ER, $r = 0.22$, S, $r = -0.09$, all P -values n.s.). We also found a positive correlation of the Mratio for ER with individual scores in EQ (Spearman $r = 0.41$, $P = 0.02$) but not for the S task (Spearman $r = -0.17$, n.s.), as well as a trend significance for a correlation of Mratio for ER with scores in IRI-EC (Spearman $r = 0.32$, $P = 0.06$), and no significant relation with the BEES for both tasks. In contrast, empathy questionnaire scores did not correlate with absolute confidence ratings or with performance accuracy (percent correct), for either ER or S task. Thus overall, several empathy measures from questionnaires were linked to metacognitive capacity in ER (A_{ROC} and Mratio), but not in S task. While these exploratory analyses indicate that individuals with high metacognitive sensitivity may have better empathic skills, they need to be further investigated and replicated in adequately powered sample size populations.

Other questionnaire scores measuring individual traits such as alexithymia (measured by TAS-20) and narcissistic personality (measured by OMNI) did not correlate with A_{ROC} values, Mratio, mean confidence rating or overall performance accuracy for each task.

Neuroimaging results

Discrimination-related activity for ER and S tasks

To examine brain activity related to perceptual processing, we contrasted the stimulus discrimination period in the ER (DISC

E) vs the S task (DISC S). For ER, we found bilateral increases in fusiform cortex ([42, -46, -23], $t = 5.80$; [-39, -46, -20], $t = 5.65$, all $P < 0.05$ FWE corrected at the peak level; Figure 3A), overlapping with face-selective areas (Kanwisher et al., 1997), as well as in the right superior temporal sulcus (STS, [51, -10, -20], $t = 5.25$, $P < 0.05$ FWE corrected at the peak level) and medial temporal structures encompassing both the left and right amygdala (but with lower cluster thresholds in the latter case; [21, -7, -26], $t = 4.25$; [-21, -10, -20], $t = 3.72$, all $P < 0.001$ uncorrected; Supplementary Table S2).

Examining the beta coefficients extracted from each peak confirmed that ER-related activation in fusiform cortex and amygdala arose only during the discrimination period, not during the confidence response period (Figure 3B, C and E). By contrast, the right STS showed an effect of task (ER vs S judgments) not only during discrimination, but also during the confidence response (Figure 3D). In addition, this region showed globally higher activity during confidence judgments relative to discrimination, unlike fusiform cortex and amygdala.

Conversely, for the S task, we found bilateral activations in the frontal eye fields (FEF, [-24, 2, 58], $t = 5.54$; [27, 5, 55], $t = 6.53$, all $P < 0.05$ FWE corrected at the peak level; Figure 3F) and parietal areas around the intraparietal sulci (IPS, [45, -34, 43], $t = 11.58$; [-45, -34, 43], $t = 6.47$, all $P < 0.05$ FWE corrected at the peak level; Supplementary Table S1), overlapping with visual-spatial processing networks (Vossel et al., 2014). Beta coefficients extracted from these regions confirmed their selective recruitment by this task during the discrimination period, not during confidence responses (Figure 3G, H, I and J).

A task (ER vs S) x judgment (DISC vs CONF) interaction contrast also corroborated specific activity increases in brain regions implicated in ER discrimination but not confidence (i.e. interaction effect [(DISC E > DISC S) > (CONF E > CONF S)]). These included the bilateral fusiform (right [42, -46, -26], $t = 4.60$, left [-39, -46, -20], $t = 5.61$, both $P < 0.05$ FWE corrected at the peak level) and occipital visual areas ([18, -97, 1], $t = 7.90$, [-18, -94, -8], $t = 9.50$, both $P < 0.05$ FWE corrected), in addition to the inferior frontal gyri (IFG, [57, 32, 4], $t = 8.65$, [-45, 20, 16], $t = 7.34$, both $P < 0.05$ FWE corrected), reflecting the role of these regions in face expression recognition (Meaux and Vuilleumier, 2016) rather than confidence.

We then tested whether brain regions recruited by the perceptual discrimination demands of each task were also involved in metacognitive sensitivity. To do so, we extracted activity parameters from specific ROIs defined by the above contrasts (Supplementary Table S2), using a sphere of 5 mm centered on peak voxels, and then correlated these activity parameters with behavioral indices of confidence (i.e. A_{ROC} values, Mratio and confidence ratings) and performance (discrimination accuracy). For ER, these perceptual ROIs included the bilateral fusiform, right STS, right amygdala and left anterior hippocampal/amygdala complex, plus the ventral and superior medial PFC (Supplementary Table S2). Correlation analyses with A_{ROC} values for ER revealed a significant modulation of the right STS ($r = .35$, $P = 0.03$) and right amygdala ($r = 0.45$, $P = 0.01$), but only the latter survived Bonferroni correction for multiple comparisons. Neither the right STS ($r = 0.02$) nor the right amygdala ($r = -0.14$) correlated with the Mratio for ER. Moreover, none of these regions correlated with A_{ROC} or Mratio values in the S task (A_{ROC} $r = 0.06$ and 0.17 , and Mratio: $r = 0.04$, $r = 0.19$ for the STS and amygdala, respectively, all P -values n.s.). All other correlations with accuracy and absolute confidence ratings were non-significant. Thus, higher activity in amygdala during the face expression discrimination phase

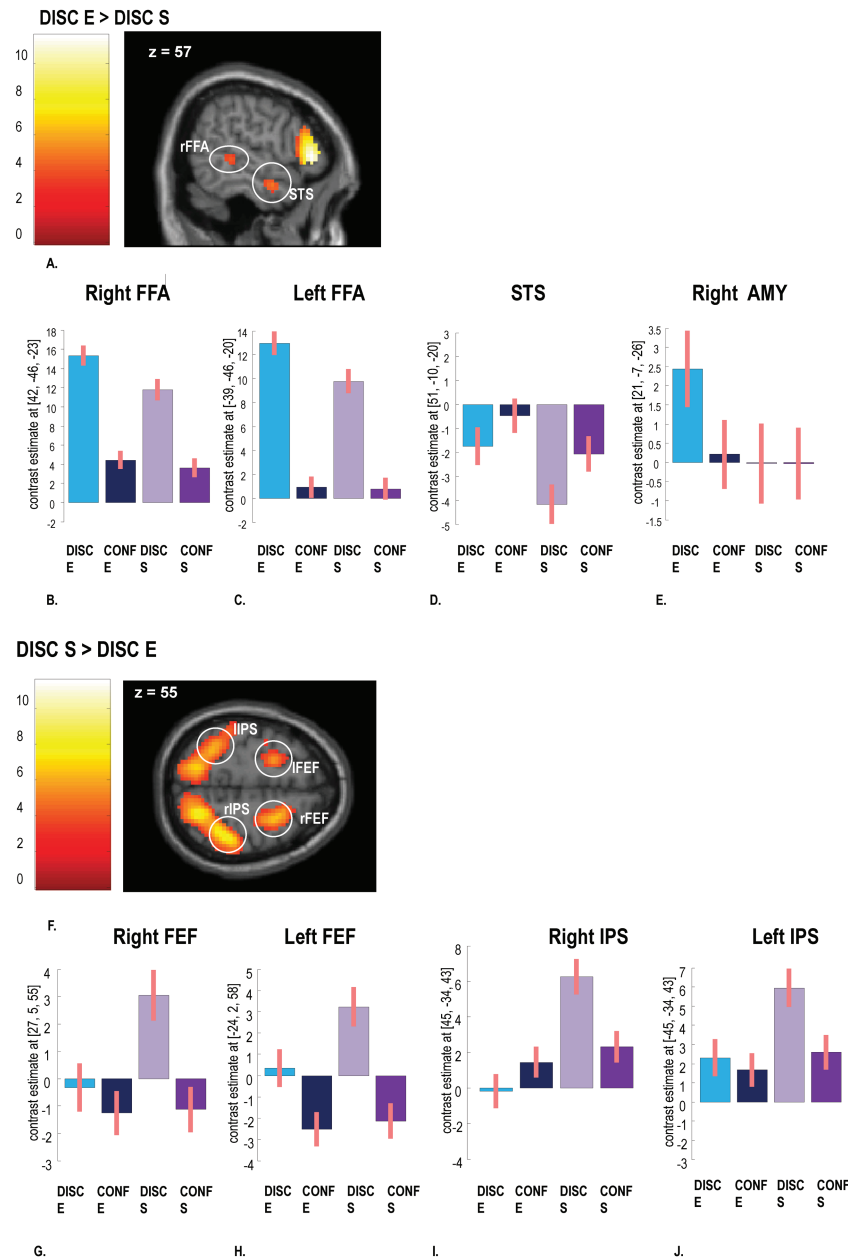


Fig. 3. fMRI results showing activations associated with stimulus discrimination. (A) Main effect of the ER task (DISC E > DISC S). Mean beta coefficients illustrating the differential effects of experimental conditions are extracted from clusters in the (B) right FFA, (C) left FFA, (D) right STS and (E) right AMY. (F) Main effect of the S task (DISC S > DISC E). Mean beta coefficients across experimental conditions from (G) right FEF, (H) left FEF, (I) right IPS and (J) left IPS. Activations maps are corrected for multiple comparisons $p_{FWE} < 0.05$ at the cluster level. Errors bars represent standard error.

predicted higher metacognitive sensitivity or, in other words, more reliable confidence evaluations; whereas activity in other face processing regions did not. However, this relationship of activity in these areas with A_{ROC} but not M_{ratio} may partly be confounded by performance effects in the ER task. Similarly, we tested whether regions involved in size discrimination correlated with metacognitive sensitivity in the latter task. Relevant ROIs included the bilateral FEF and IPS (Supplementary Table S2). None of these correlations reached significance. These results argue against a direct link between metacognitive sensitivity and brain areas directly supporting visual discrimination processes, and suggest mechanistic differences between metacognition for ER and S performance.

Confidence-related activity

Next, we examined brain regions specifically engaged during confidence monitoring. To identify confidence processes related to each task, we compared response periods corresponding to the confidence judgment following ER (CONF E > CONF S) and S (CONF S > CONF E). Confidence monitoring for ER produced strong activation in the retrosplenial cortex/posterior cingulate (RSC/PCC, left peak [-18 -61, 16], $t = 4.70$, $P < 0.05$ FWE corrected at the peak and cluster level, extending to the right [12, -64, 19], $t = 3.84$, $P < 0.001$ uncorrected). A smaller cluster was also found in the right STS ([51, -10, -20], $t = 3.93$, $P < 0.001$ uncorrected), overlapping with the region activated during the ER discrimination period (Supplementary Table S2). For the opposite

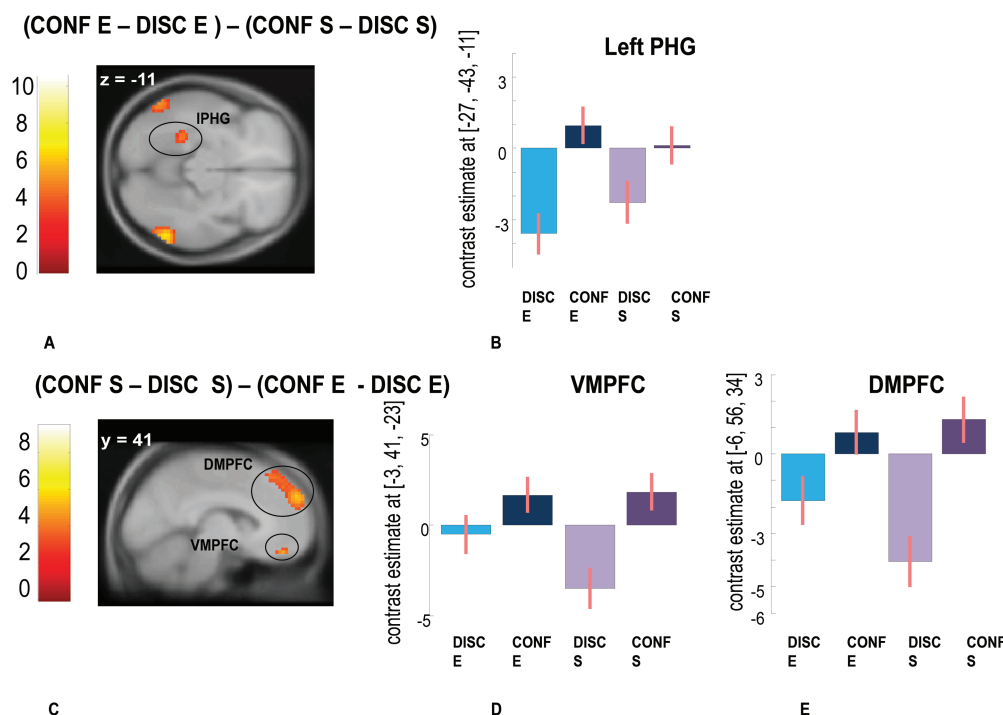


Fig. 4. fMRI results showing activations associated with confidence. (A) Interaction contrast testing for greater confidence effects during the ER than during S task [(CONF E > DISC E) > (CONF S > DISC S)], showing activation in the IPHG with (G) corresponding beta coefficients. (B) Activation of the DMPFC and VMPFC observed in the opposite interaction contrast [(CONF S > DISC S) > (CONF E > DISC E)] testing for greater confidence effects during S than ER recognition, with (C, D) corresponding beta coefficients. Activations maps corrected for multiple comparisons pFWE < 0.05 at the cluster level. Errors bars represent standard error.

comparison (CONF S > CONF E), no brain region reached statistical significance (at $P < 0.00$ uncorrected).

To confirm the role of distinct networks for metacognition of ER and S judgments, we also computed a direct interaction contrasts (CONF E > DISC E) > (CONF S > DISC S) and the reverse (CONF S > DISC S) > (CONF E > DISC E), testing for specific increases during confidence but not discrimination judgments in each task (Supplementary Table S1). The former confirmed increased activity in both the right ([15, -58, 22], $t = 4.16$) and the left RSC/PCC ([-12, -61, 22], $t = 4.72$, $P < 0.05$ FWE corrected at the cluster level). Increased brain activity in the right RSC/PCC extended superiorly to the precuneus on the right ([6, -37, 46], $t = 4.14$, $P < 0.05$ FWE corrected at the cluster level) and left side ([-6, -37, 46]; $t = 4.08$, $P < 0.05$ FWE corrected at the cluster level) as well as the left parahippocampal gyrus (IPHG) [-27, -43, -11], $T = 4.41$, $P < 0.05$ FWE corrected at the cluster level (Figure 4A and B). These regions thus showed selective engagement by metacognitive confidence monitoring (but not discrimination) during face expression recognition.

The opposite contrast testing for differential involvement in confidence during the S task (CONF S > DISC S) > (CONF E > DISC E) revealed activations in ventromedial prefrontal cortex (VMPFC, [-3, 41, -23], $t = 5.11$, $P < 0.05$ FWE corrected at the peak level) and dorsomedial prefrontal cortex (DMPFC, [-6, 56, 34], $t = 5.32$, $P < 0.05$ FWE corrected at the peak and cluster level, due to an increase during confidence judgments paralleled by a decrease during discrimination judgments in this task (Figure 4C, D and E).

To examine whether regions engaged during confidence judgments were modulated in direct proportion with increasing metacognitive sensitivity, we extracted activity parameters

(beta values) from specific ROIs defined by the above contrasts (CONF E > DISC E) > (CONF S > DISC S) (including bilateral RSC/PCC, bilateral precuneus and IPHG) with a sphere of 5 mm centered on peak voxels, and correlated them with individual measures of metacognitive abilities (A_{ROC} and M_{ratio} values). For ER metacognition, this analysis revealed a significant correlation in the right PCC/RSC with M_{ratio} ([15, -58, 22], $r = -0.36$, $P = 0.02$) but not A_{ROC} ($r = -0.08$, n.s.). All other correlations with M_{ratio} and A_{ROC} for S and ER in these regions did not pass the statistical threshold.

We repeated the same analysis as above and correlated beta values in ROI engaged by confidence in the S task, using a sphere of 5 mm centered in peak voxels of VMPFC and DMPFC issued from (CONF S > DISC S) > (CONF E > DISC E). None showed any correlation with A_{ROC} and M_{ratio} for ER and S task (ER task A_{ROC} Spearman $r = -0.20$, M_{ratio} $r = 0.03$; S task A_{ROC} Spearman $r = -0.20$, M_{ratio} Spearman $r = -0.03$, all P -values n.s.).

Finally, we computed a conjunction analysis, i.e. [(CONF S > DISC S) AND (CONF E > DISC E)] to determine whether metacognitive processes also recruited some common brain regions. This analysis showed activation in motor and somatosensory regions (including left primary motor cortex, bilateral somatosensory regions, supplementary motor area and posterior putamen), reflecting motor response execution shared across tasks (key presses), but also in bilateral rostromedial prefrontal cortex (medial BA 10), orbitofrontal (OFC) and VMPFC, plus the midcingulate cortex (extending to the precuneus) and the IPHG (see Figure 5 and Supplementary Table S3 for full details). The latter set of areas points to a metacognitive network involved in post-decision confidence processes across different task domains.

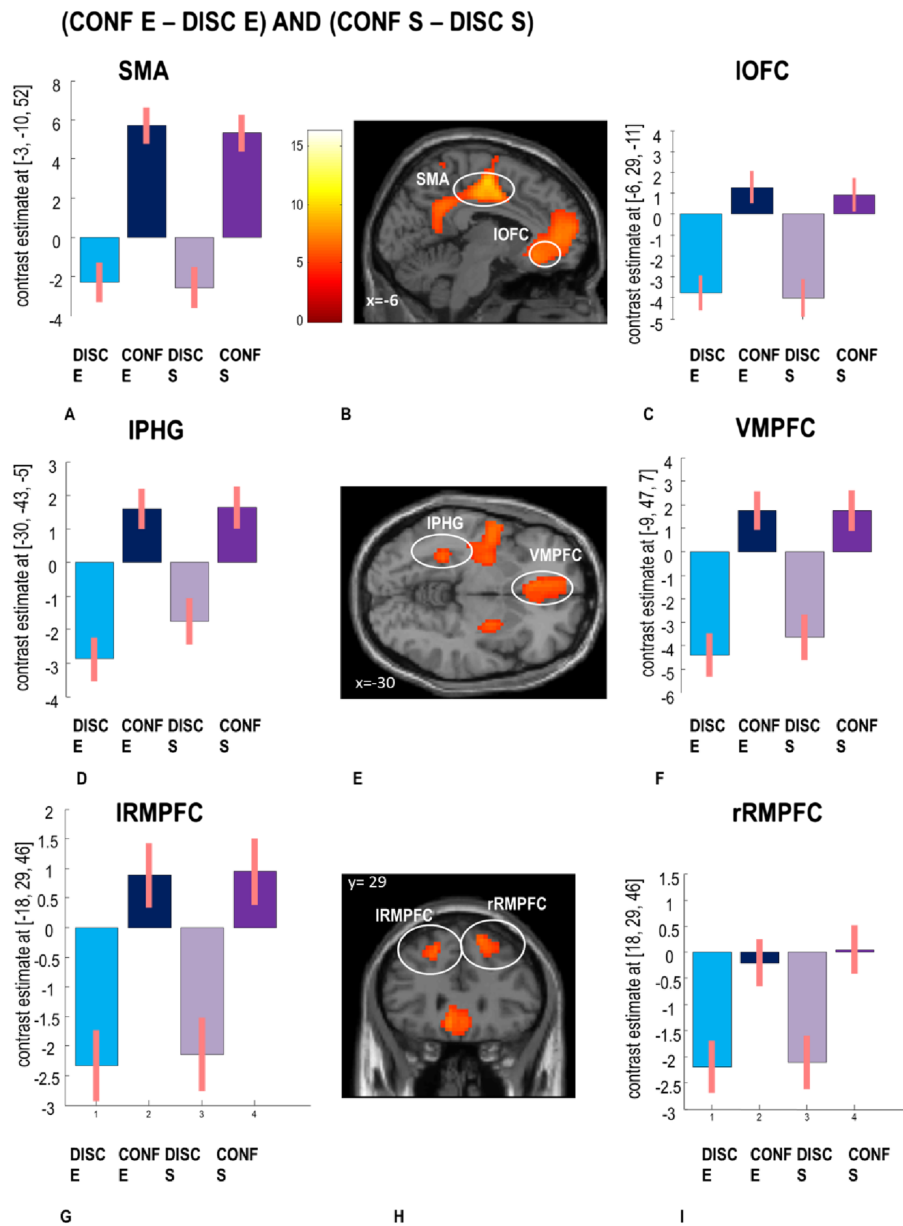


Fig. 5. fMRI results showing common regions for confidence across tasks (conjunction analysis). Activations are shown for (B) the SMA and OFC, with (A, C) corresponding beta coefficients, as well as for (E, D, F) the IPHG and VMPFC, and (H, G, I) the left and right rostromedial PFC (rRMPPFC), as identified in the conjunction contrast [(CONF S > DISC S) AND (CONF E > DISC E)] testing for common confidence effects across both tasks. Activations maps corrected for multiple comparisons $p_{FWE} < 0.05$ at the cluster level. Errors bars represent standard error.

Parametric analysis of confidence

For completeness, we tested for brain regions displaying any confidence-related activity that increased with increasing metacognitive sensitivity using a separate parametric whole-brain regression analysis. To do so, brain activity during confidence responses for ER (CONF E > CONF S) and S tasks (CONF S > CONF E) was correlated with individual measures of metacognitive sensitivity (A_{ROC} values, M_{ratio}) as linear parametric factors. For ER, results revealed a significant modulation by A_{ROC} in the PCC and RSC extending rostrally to the midcingulate cortex, but also IPHG and posterior parietal areas (Supplementary Table S4). Interestingly, while activations in PCC overlapped with areas specifically engaged during confidence judgments in ER (Supplementary Table S2), those

in the midcingulate/precuneus and PHG partly overlapped with areas shared by confidence processes across both tasks (Supplementary Table S4).

Regression analysis with the same contrast (i.e. CONF E > CONF S) but using A_{ROC} values from the S task revealed no significant correlation. Moreover, no significant correlation was found between confidence-related activity in the S task (CONF S > CONF E) and A_{ROC} or M_{ratio} values for either ER or S conditions.

Additional analyses also revealed that PCC activity was neither modulated by absolute confidence levels for ER judgments nor by task performance (percent correct or d_{prime}). Thus, activation in the PCC reflected the degree of metacognitive sensitivity, not overall confidence *per se*. None of these regression

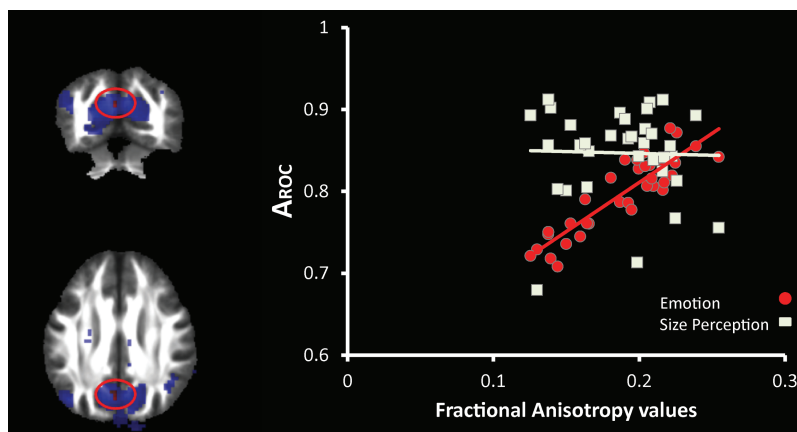


Fig. 6. DTI results. Statistical maps of the whole-brain correlation between white matter structure (FA values) and metacognition for ER (A_{ROC} values), illustrating the correspondence with functional MRI results. (A) A significant correlation between FA and A_{ROC} values for ER was observed in overlapping voxels (in red, cluster peak at $[-3, -67, 28]$, $P < 0.05$ corrected for multiple comparisons after small volume correction based on the functional mask; circled by a red ellipse on the coronal and axial views). Functional cluster (in blue) activated in the contrast $CONF E > CONF S$, thresholded at $P < 0.05$, uncorrected for illustration purpose. (B) Scatter plot of FA values at $[-3, -67, 28]$ against A_{ROC} values for ER (in red) and S (in white) tasks.

analyses showed significant correlations with individual scores in empathy questionnaires.

Diffusion tensor imaging results

Finally, to determine whether differences in metacognitive sensitivity and empathy skills also implied differences in structural connectivity between brain areas, we tested for any relation between indices of metacognition (A_{ROC}) in both tasks, and measures of fiber density in white matter tracts using DTI. A whole-brain correlation between FA and A_{ROC} values from the ER task revealed a unique cluster of significant positive correlation ($P < 0.02$ corrected for multiple comparisons, 11 voxels) in the white matter beneath the RSC/PCC (MNI coordinates $[-3, -67, 28]$, Figure 6A and B). The latter region neatly corresponded to the RSC/PCC area that was functionally modulated by confidence judgments ($CONF E > CONF S$), as independently identified in the whole-brain fMRI analysis above. Moreover, the same cluster also showed a trend for a correlation with empathy skills measured by BEES scores (Spearman $r = 0.34$, $P = 0.054$). Correlations with IRI scores or other personality trait questionnaires' scores were not significant. Correlation analysis of DTI data with A_{ROC} values from the S task showed no significant results.

Discussion

In our study, participants had to evaluate confidence of their accuracy in either face expression recognition or visual size discrimination. Behavioral performance (percent correct) reliably correlated with stimulus intensity in both tasks (i.e. distance between morphed expressions for ER and difference in line thickness for S judgments), demonstrating that participants extracted relevant visual information in each task while the most ambiguous stimuli led to chance-level performance (Figure 2D and F). Confidence of visual perceptual discrimination judgments (S task) also correlated positively with stimulus intensity, reflecting greater confidence for more perceptually distinct trials (Figure 2G). By contrast, confidence of ER did not show such relation to expression intensity (Figure 2B). Even for highly ambiguous stimuli that could not be easily classified as

angry or happy, participants reported high confidence of their ER ability. This effect was reflected in the mean A_{ROC} (Figure 2B), showing worse metacognition in face expression recognition than visual perception, in addition to mean confidence ratings (Figure 2C) and OU/C scores showing overconfidence for ER. Moreover, individual metacognitive abilities for visual perception did not predict those for ER. However, the efficiency of metacognitive mechanisms (indexed by the M_{ratio}) was similar for both tasks, suggesting that task-related information was equally available for metacognitive judgments in both domains. It is important to note that A_{ROC} and M_{ratio} are obtained by calculations attempting to relate confidence to performance in order to infer metacognitive ability, but none of these measures considers stimulus intensity. Here we found that even faces with ambiguous expressions evoked similar confidence ratings as faces with more intense expressions. In other words, one cannot predict participants' accuracy based on their expressed levels of confidence for ER, unlike for visual size perception (where most ambiguous stimuli elicit both lower accuracy and lower confidence; see Figure 2D, E, F and G).

Altogether, these behavioral findings converge to suggest that at least partly distinct mechanisms subserved metacognition for ER and size perception. A possible limitation to our conclusion is that discrimination performance was slightly but significantly lower in the ER than S task (79% vs 82% correct), despite our staircase procedure. However, again, whereas subjects performed the primary discrimination tasks psychometrically in both cases, their confidence for ER (but not S) started off really high, even for the more ambiguous faces, suggesting a general overestimation of their capacities for emotional recognition. By contrast, in the S task, confidence followed a more progressive psychometric function shape. Moreover, strikingly, confidence of ER decisions was higher despite lower performance.

Inflated confidence and worse metacognition for ER accord with psychology literature reporting that people tend to overestimate their empathic accuracy (i.e. correctly inferring thoughts and feelings of other people) (Ickes, 2003; Ickes et al., 1990; Marangoni, Garcia et al., 1995), as well as their ability to detect emotions in pictures of faces and speech (Realo et al., 2003). Poor calibration of confidence for ER might result from the lack of feedback in real life (e.g. due to social norms) as

well as insufficient information or deceptive intentions from the expresser (e.g. when we ask others if they are angry or sad and they say 'I'm ok'). Alternatively, overconfidence of ER may result from a data sampling bias during the face expression categorization process, possibly because the latter unfolds quickly and automatically (80–100 ms post-stimulus onset) (Meaux et al., 2013) based on partial cues from particular features, such as the eye or mouth region (Eisenbarth and Alpers, 2011; Smith et al., 2005). Even though rapid detection of facial emotion is crucial for efficient behavior, in the case of ambiguous situations it has been shown that preferential focus on the strength of a limited source of information with insufficient regard for its weight relative to more general cues or knowledge can contribute to erroneous conclusions of high confidence (Griffin and Tversky, 1992; Zylberberg et al., 2014). By analogy, focusing on single face parts with insufficient use of global cues from the whole face expression configuration (Meaux and Vuilleumier, 2016) might be an important factor promoting overconfidence. Moreover, both explanations of incomplete feedback and sampling biases are not mutually exclusive. Such biases in metacognition due to absent feedback or reliance on partial cues would not affect more automatic mechanisms subserving face perception, which typically operate without conscious deliberative processing.

In keeping with our behavioral findings, we found strong activations in distinct brain networks engaged by confidence processes for each decision type, adding support to a domain-specific view proposed by other studies (Baird et al., 2013; McCurdy et al., 2013). Metacognition for ER recruited the RSC/PCC, extending to the precuneus and IPHG, whereas metacognition for visual perception engaged DMPFC and VMPFC. Remarkably, DTI results converged with fMRI data by further highlighting that structural white matter connectivity of the RSC/PCC and precuneus was associated with higher metacognitive sensitivity for ER (but not S). In addition, we found that PCC/RSC activity during confidence judgments for ER (but not S) correlated with metacognitive efficiency for ER (indexed by the *Mratio*), although it did not predict higher confidence (i.e. absolute levels) *per se*. Together, these findings provide strong evidence that these regions play a unique role in monitoring ER abilities, presumably by integrating signals from remote brain areas during or after the decision phase.

The RSC/PCC complex is particularly responsive to contextual and memory information associated with emotional significance (Hofstetter et al., 2012). It is among the most consistently activated regions in response to emotionally salient stimuli, possibly reflecting prior associations stored in memory (Maddock et al., 1999). It also belongs to a distributed network that mediates self-referential introspection and recollection (Kim, 2012; Schmitz and Johnson, 2007; Whitfield-Gabrieli et al., 2011), representation of self-knowledge in time (D'Argembeau and Salmon, 2012), and decision-making based on internal preferences under conditions where there is no objectively correct response (Johnson et al., 2005). High confidence during memory recall (Chua et al., 2009a) is associated with increased activity in RSC/PCC (Kim and Cabeza, 2009). Gray matter volume in the RSC/PCC also correlates with metacognitive sensitivity (indexed by the *Mratio*) during visuomotor decisions about reaching movements (Sinanaj et al., 2015). Therefore, our findings that the RSC/PCC is functionally activated by confidence judgment during ER, and that its structural connectivity predicts the reliability of confidence relative to objective performance, provide novel evidence that metacognition for ER requires the integration of visual information with internal priors that can be shaped by contextual

information, memory associations, past knowledge and other social and self-relevant cues. Integration of prior mnemonic information with current metacognitive experiences may be crucial to applying general theory-based metacognitive knowledge ('I am good with faces') to online experience-based metacognition ('I am sure she is happy'). We surmise that the RSC/PCC constitutes a crucial hub for such integration, allowing adequate weighting of sensory evidence by relevant priors to determine self-confidence in ER (Griffin and Tversky, 1992). This interpretation brings empirical support to recent proposals that metacognition may involve evaluative processes beyond purely perceptual decision mechanisms, including memory-based information (Shadlen and Shohamy, 2016). Mnemonic contextual contributions subserved by the RSC/PCC may be more important for emotional recognition than perceptual decision in line with a key role of metacognition in suprapersonal control (Shea et al., 2014), and the fact that veridical feedback about emotions felt by other is often unavailable or unreliable in many real-life situations (Ekman, 2003; Ekman et al., 1988). However, we cannot exclude that these differences in brain activity may also partly be linked to differential computations to estimate confidence as perceived uncertainty in the estimated variable (Navajas et al., 2017; Pouget et al., 2016).

In agreement with a role of memory processes, the IPHG was also modulated by confidence judgments during ER, as shown in the task \times judgment interaction analysis (Figure 4F and G). Again, PHG is reported to activate during internally-oriented cognition, online self-monitoring and introspective mental activity (Chua et al., 2006; Ward et al., 2014). Recent studies found PHG activation during retrospective confidence evaluation in a face-name associative memory task, particularly during high-confidence trials (Chua et al., 2006; Moritz et al., 2006). Moreover, grey matter volume in both the PHG and precuneus correlates with reality monitoring performance (Buda et al., 2011). Precuneus has also been implicated in metamemory (Baird et al., 2013; McCurdy et al., 2013) and visuomotor metacognition (Sinanaj et al., 2015). Interestingly, however, we found that both the PHG and precuneus were among regions commonly recruited during confidence judgments in our two tasks (see conjunction analysis, Figure 5 and Supplementary Table S3). Taken together these findings suggest that metacognition is supported by both common and specific neural systems across different cognitive domains. Thus, while the RSC/PCC may be distinctively implicated in metacognition for ER, the PHG and precuneus may contribute to different domains of metacognition (Morales et al., 2018), presumably reflecting memory associations and self-referential information, which might be more important for ER than visual size judgments in the present study. Interestingly, we also recently showed selective activation in PHG in patients suffering from functional motor weakness (conversion disorder, i.e. patients who maintain they are paralyzed despite intact neuroanatomical pathways), relative to healthy controls, when they report metacognitive confidence during a visuomotor task (Bègue et al., 2018). Altogether, these data converge to support the notion that memory and experience-based processes may contribute to metacognitive functions, at least in certain domains or conditions.

In contrast to ER recognition, confidence for visual perception engaged the DMPFC and VMPFC. This confirms previous work on non-emotional perceptual tasks (Fleming et al., 2012) where DMPFC was activated when participants reported their confidence during a house-face discrimination paradigm. VMPFC activity has been linked with prospective metamemory (Chua et al., 2009b; Kao et al., 2005; Pannu and Kaszniak, 2005), and

subjective confidence in value-based choices (De Martino et al., 2013; Lebreton et al., 2015). Our results thus accord with a role of these regions in metacognition (Fleming and Dolan, 2012; Fleming et al., 2012). Additional regions within a distributed network were also commonly activated during confidence judgments in both tasks (but not during perceptual decisions themselves), involving dorsolateral prefrontal and posterior parietal cortices (PPC), as well as orbitofrontal and anterior cingulate regions bilaterally (Figure 5 and Supplementary Table S3) in line with previous neuroimaging literature on metacognition (Fleming and Dolan, 2012). This network could underpin various processes related to motor preparation, attention and working memory (Dorsolateral prefrontal cortex and PPC), as well as internal valuation (VMPFC and OFC) and error monitoring (Supplementary motor area (SMA), shared between tasks at the time of confidence ratings).

Altogether, these findings unveil both distinct and common mechanisms for metacognition of ER and visual perception, with a crucial role of memory systems during ER in particular. These systems, centered on the RSC/PCC, but also the PHG and precuneus, are well positioned to mediate post-decisional evaluative processes integrating sensory evidence with internal priors and knowledge. This adds novel support to recent theoretical accounts of metacognition underscoring a role of memory-based processes (Shadlen and Shohamy, 2016). Future studies should clarify whether individual differences in these memory systems (e.g. variability in RSC/PCC activity and anatomy predicting better ER metacognition) are inborn or shaped by experience.

As anticipated, judging emotional expressions strongly activated the bilateral fusiform cortex, STS and amygdala, known to be respectively implicated in face recognition (Kanwisher and Yovel, 2006), gaze and expression processing (Engell and Haxby, 2007) and emotional learning (Adolphs et al., 1994; N'Diaye et al., 2009). Instead, judging visual size activated frontal parietal regions implicated in visuospatial processing and eye movements (Burman and Segraves, 1994; Vernet et al., 2014). However, with the exception of the right STS, none of these areas showed differential activity between the two tasks during confidence judgments. Moreover, although these areas directly supported face expression or size discrimination, respectively, their activation at the time of the discrimination generally did not predict confidence levels or metacognition sensitivity. Only activity in the right amygdala correlated with better metacognition for ER (with A_{ROC} , but not with other performance measures). This finding is consistent with a role of the amygdala in emotional face recognition (Adolphs et al., 1994; Cristinzio et al., 2010) and accords with a Bayesian framework asserting that confidence is computed using neural signals elicited or accumulated during task-related decisions (Pouget et al., 2016).

In keeping with the latter view, compelling evidence from animal research suggests that regions implicated in stimulus discrimination carry information that is used for metacognitive evaluation, presumably reflecting a common variable that gives rise to both decision and confidence in the decision (Fetsch et al., 2014; Kiani and Shadlen, 2009; Moran et al., 2015; Pleskac and Busemeyer, 2010; Ratcliff and Starns, 2013). In humans who perform a perceptual categorization task on ambiguous stimuli (house vs face), both the amygdala and STS are modulated by higher levels of sensory evidence, suggesting a role in evidence accumulation prior to perceptual decision (Filimon et al., 2013). However whether these decision-related structures are able to read-out confidence information remains unclear (Grimaldi et al., 2015). Thus, our results leave open the question as to whether, at the time of ER decisions, the amygdala computes

the available sensory information for confidence estimation, or whether confidence-related activity in the amygdala is relayed to other brain regions (e.g. in prefrontal areas or RSC/PCC). Hence, although our correlational results about amygdala activity highlight a complex relationship between perceptual discrimination mechanisms and confidence, they do not allow teasing apart a role for decisional or post-decisional processes in metacognition. Intriguingly, these results also indicate that, at least for ER recognition, different decision-related brain regions may make distinct contributions to confidence judgments, possibly intervening at different stages of this process (e.g. at decision time only, e.g. amygdala; or during both decision and confidence, e.g. STS). Other regions in the face processing network may primarily subservise perceptual decisions with no impact on confidence evaluation (e.g. fusiform) or vice versa (RSC/PCC). More generally, our new data also unveil an important role of human amygdala in mediating reliable confidence of ER recognition, possibly via accurate ER, in accordance with its major role in social appraisal and social learning

Finally, in line with previous behavioral work (Kelly and Metcalfe, 2011), we found that confidence of ER by itself did not correlate with empathy scores. That is, people who claim higher confidence in their ER capacity have no higher empathy than those who are less confident. Likewise, there was no correlation between empathy and discrimination accuracy in the ER task. In contrast, however, high metacognition for ER (indexed by AROC) was predictive of higher emotional empathy scores as indexed by the IRI empathic concern subscale (Davis, 1980) measuring emotional sharing and sympathizing with others, and by the BEES (Mehrabian, 1996) measuring vicarious experiences of the emotional feelings of others. Higher metacognition efficiency for ER (indexed by Mratio) also correlated significantly with the EQ of Baron-Cohen (Baron-Cohen and Wheelwright, 2004) and at trend level with IRI-EC (Davis, 1980). These preliminary findings suggest that the capacity to empathize with others may partly rely on accurate judgment of one's own capacity in ER recognition. Different authors have argued for a close functional relationship between empathy and metacognition (Caruthers, 2009; Frith, 2012). Moreover, deficits in metacognition are thought to contribute to failures in social cognition (and vice versa) in several neuropsychiatric disorders (David et al., 2012). Our current results provide novel preliminary evidence in favor of a relationship between emotional empathy and metacognition for ER, supporting the notion of partly shared rather than distinct processes in line with recent work (Valk et al., 2016). Future studies should be designed to further investigate these relationships in depth.

Strengths and limitations

Our experimental design allowed us not only to directly compare two domains of metacognition in the same paradigm but also to isolate processes linked to perceptual decision and confidence for both tasks by using a time jitter between these two judgments. For both tasks, we found dissociable patterns of brain response to perceptual decision and confidence, converging with the view that metacognitive experiences do not simply emerge from sensory evidence accumulated during the task at hand (e.g. see Barron et al., 2013; Flavell, 1979; Wimmer and Shohamy, 2012). A direct demonstration of such dissociation has not been clearly established up to now (see Shadlen and Shohamy, 2016).

However, a few limitations need to be considered. The observed differences in performance between ER and S may be an indication that a higher number of trials is needed

to stabilize ER performance compared to S performance via staircase, and future studies should consider these domain-specific characteristics. Our exploratory results concerning interindividual differences should also be considered as preliminary given our relatively modest sample size for reliable correlation analyses, and will therefore need to be further investigated with larger adequately powered cohorts. Finally, higher confidence ratings for ER might suggest ceiling effects arising exclusively in this task. However, ceiling effect or rating biases are unlikely given that participants used the same range of rating values in both tasks and that performance was globally lower in the ER task.

Conclusion

Our study is the first to unveil mechanisms underlying metacognitive confidence in ER. Behaviorally, we report relatively worse metacognition for ER as compared with basic visual perceptual decisions in that (i) confidence tended to be high independent of stimulus intensity levels and (ii) inflated subjective trust in performance was inflated (indexed by the O/UC score). However, it is worth noting that the efficiency of metacognitive mechanisms weighted by task performance (as indexed by the Mratio) is similar for both our tasks. Moreover, we find that metacognitive ability in ER (indexed by AROC and Mratio, but not good recognition *per se*) positively correlates with individual empathy skills.

At the brain level, our fMRI and DTI results converge to demonstrate a key role for limbic regions in metacognition for ER, including the RSC/PCC that may subserve integration with stored associations and self-reflective processes allowing metacognitive monitoring of recognition accuracy. Only the right amygdala showed activation during facial expression recognition that predicted metacognitive sensitivity, in line with its role in processing emotional faces. More generally, our study adds novel support to the emerging notion that integrative and memory-based mechanisms may serve to weight sensory evidence during confidence judgments (Shadlen and Shohamy, 2016).

Funding

This Center of Competence in Research (NCCR) Affective Sciences, financed by the Swiss National Science Foundation (51NF40-104897) and hosted by the University of Geneva, and by Foremane Fund award (to P.V.).

Author contributions

I.B., S.S. and P.V. designed research. I.B., J.H. and M.P. acquired data. I.B. and M.V. analyzed data. I.B., M.V., S.S. and P.V. wrote and reviewed versions of the paper.

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

Adams, J.K., Adams, P.A. (1961). Realism of confidence judgments. *Psychological Review*, **68**, 33–45 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/13681400>.

- Adolphs, R., Tranel, D., Damasio, H., Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, **372**(6507), 669–72. doi:10.1038/372669a0.
- Bagby, R.M., Taylor, G.J., Parker, J.D. (1994). The twenty-item Toronto Alexithymia Scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, **38**(1), 33–40.
- Baird, B., Smallwood, J., Gorgolewski, K.J., Margulies, D.S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *Journal of Neuroscience*, **33**(42), 16657–65. doi:10.1523/JNEUROSCI.0786-13.2013.
- Baron-Cohen, S., Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, **34**(2), 163–75.
- Barron, H.C., Dolan, R.J., Behrens, T.E.J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience*, **16**(10), 1492.
- Begue, I., Blakemore, R., Klug, J., et al. (2018). Metacognition of visuomotor decisions in conversion disorder. *Neuropsychologia*, **114**, 251–65. doi:10.1016/j.neuropsychologia.2018.04.018.
- Buda, M., Fornito, A., Bergstrom, Z.M., Simons, J.S. (2011). A specific brain structural basis for individual differences in reality monitoring. *Journal of Neuroscience*, **31**(40), 14308–13. doi:10.1523/JNEUROSCI.3595-11.2011.
- Burman, D.D., Segraves, M.A. (1994). Primate frontal eye field activity during natural scanning eye movements. *Journal of Neurophysiology*, **71**(3), 1266–71.
- Carruthers, P. (2009). Mindreading underlies metacognition. *Behavioral and Brain Sciences*, **32**(2), 164–82. doi:10.1017/S0140525x09000545.
- Chua, E.F., Schacter, D.L., Rand-Giovannetti, E., Sperling, R.A. (2006). Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. *NeuroImage*, **29**(4), 1150–60. doi:10.1016/j.neuroimage.2005.09.058.
- Chua, E.F., Schacter, D.L., Sperling, R.A. (2009a). Neural basis for recognition confidence in younger and older adults. *Psychology and Aging*, **24**(1), 139–53. doi:10.1037/A0014029.
- Chua, E.F., Schacter, D.L., Sperling, R.A. (2009b). Neural correlates of metamemory: a comparison of feeling-of-knowing and retrospective confidence judgments. *Journal of Cognitive Neuroscience*, **21**(9), 1751–65. doi:10.1162/jocn.2009.21123.
- Cristinzio, C., N'Diaye, K., Seeck, M., Vuilleumier, P., Sander, D. (2010). Integration of gaze direction and facial expression in patients with unilateral amygdala damage. *Brain*, **133**, 248–61. doi:10.1093/brain/awp255.
- D'Argembeau, A., Salmon, E. (2012). The neural basis of semantic and episodic forms of self-knowledge: insights from functional neuroimaging. *Advances in Experimental Medicine and Biology*, **739**, 276–90. doi:10.1007/978-1-4614-1704-0_18.
- David, A.S., Bedford, N., Wiffen, B., Gilleen, J. (2012). Failures of metacognition and lack of insight in neuropsychiatric disorders. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, **367**(1594), 1379–90.
- Davis, M.H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach, *Journal of personality and social psychology* **44**(1), 113.
- De Martino, B., Fleming, S.M., Garrett, N., Dolan, R.J. (2013). Confidence in value-based choice. *Nature Neuroscience*, **16**(1), 105–10. doi:10.1038/nn.3279.

- Eisenbarth, H., Alpers, G.W. (2011). Happy mouth and sad eyes: scanning emotional facial expressions. *Emotion*, *11*(4), 860–65. doi:10.1037/a0022758.
- Ekman, P. (2003). Darwin, deception, and facial expression. *Annals of the New York Academy of Science*, *1000*, 205–21 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14766633>.
- Ekman, P., Friesen, W.V., Osullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology*, *54*(3), 414–20. doi:10.1037//0022-3514.54.3.414.
- Engell, A.D., Haxby, J.V. (2007). Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia*, *45*(14), 3234–41. doi:10.1016/j.neuropsychologia.2007.06.022.
- Fetsch, C.R., Kiani, R., Newsome, W.T., Shadlen, M.N. (2014). Effects of cortical microstimulation on confidence in a perceptual decision. *Neuron*, *83*(4), 797–804. doi:10.1016/j.neuron.2014.07.011.
- Filimon, F., Philastides, M.G., Nelson, J.D., Kloosterman, N.A., Heekeren, H.R. (2013). How embodied is perceptual decision making? Evidence for separate processing of perceptual and motor decisions. *Journal of Neuroscience*, *33*(5), 2121–36. doi:10.1523/Jneurosci.2334-12.2013.
- Fischhoff, B., Slovic, P., Lichtenstein, S. (1977). Knowing with certainty—appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(4), 552–64. doi:10.1037//0096-1523.3.4.552.
- Flavell, J.H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.) *The Nature of Intelligence*, (pp. 231–235). Hillsdale, NJ: Lawrence Erlbaum.
- Flavell, J.H. (1979). Metacognition and cognitive monitoring: a new area of cognitive—developmental inquiry. *American Psychologist*, *34*(10), 906.
- Fleming, S.M., Dolan, R.J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *367*(1594), 1338–49. doi:10.1098/rstb.2011.0417.
- Fleming, S.M., Huijgen, J., Dolan, R.J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *Journal of Neuroscience*, *32*(18), 6117–25. doi:10.1523/JNEUROSCI.6489-11.2012.
- Fleming, S.M., Lau, H.C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, p. 443.
- Fleming, S.M., Weil, R.S., Nagy, Z., Dolan, R.J., Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541–3. doi:10.1126/science.1191883.
- Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., Frackowiak, R.S.J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *3*(3), 165–89. doi:10.1002/Hbm.460030303.
- Frith, C.D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society London Series B: Biological Sciences*, *367*(1599), 2213–23. doi:10.1098/rstb.2012.0123.
- Galvin, S.J., Podd, J.V., Drga, V., Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions. *Psychonomic Bulletin and Review*, *10*(4), 843–876 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15000533>.
- Griffin, D., Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*(3), 411–35. doi:10.1016/0010-0285(92)90013-R.
- Grimaldi, P., Lau, H., Basso, M.A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: confidence in decision-making. *Neuroscience and Biobehavioral Reviews*, *55*, 88–97. doi:10.1016/j.neubiorev.2015.04.006.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–33. doi:10.1016/S1364-6613(00)01482-0.
- Hofstetter, C., Achaibou, A., Vuilleumier, P. (2012). Reactivation of visual cortex during memory retrieval: content specificity and emotional modulation. *NeuroImage*, *60*(3), 1734–45. doi:10.1016/j.neuroimage.2012.01.110.
- Ickes, W. (2003). *Everyday Mind Reading: Understanding What Other People Think and Feel*, Amherst, NY: Prometheus Books.
- Ickes, W., Stinson, L., Bissonnette, V., Garcia, S. (1990). Naturalistic social cognition—empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, *59*(4), 730–42. doi:10.1037//0022-3514.59.4.730.
- Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M. (2012). FSL. *NeuroImage*, *62*(2), 782–90. doi:10.1016/j.neuroimage.2011.09.015.
- Johnson, S.C., Schmitz, T.W., Kawahara-Baccus, T.N., et al. (2005). The cerebral response during subjective choice with and without self-reference. *Journal of Cognitive Neuroscience*, *17*(12), 1897–906. doi:10.1162/089892905775008607.
- Jones, D.K. (2004). The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. *Magnetic Resonance in Medicine*, *51*(4), 807–15. doi:10.1002/Mrm.20033.
- Josephs, O., Henson, R.N. (1999). Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *354*(1387), 1215–28. doi:10.1098/rstb.1999.0475.
- Kanwisher, N., McDermott, J., Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–11.
- Kanwisher, N., Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–28. doi:10.1098/rstb.2006.1934.
- Kao, Y.C., Davis, E.S., Gabrieli, J.D. (2005). Neural correlates of actual and predicted memory formation. *Nature Neuroscience*, *8*(12), 1776–83. doi:10.1038/nn1595.
- Kelly, K.J., Metcalfe, J. (2011). Metacognition of emotional face recognition. *Emotion*, *11*(4), 896.
- Kepecs, A., Uchida, N., Zariwala, H.A., Mainen, Z.F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–31. doi:10.1038/nature07200.
- Kiani, R., Shadlen, M.N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*(5928), 759–64. doi:10.1126/science.1169405.
- Kim, H. (2012). A dual-subsystem model of the brain's default network: self-referential processing, memory retrieval processes, and autobiographical memory retrieval. *NeuroImage*, *61*(4), 966–77. doi:10.1016/j.neuroimage.2012.03.025.
- Kim, H., Cabeza, R. (2009). Common and specific brain regions in high- versus low-confidence recognition memory. *Brain Research*, *1282*, 103–13. doi:10.1016/j.brainres.2009.05.080.
- Lebreton, M., Abitbol, R., Daunizeau, J., Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience*, *18*(8), 1159–67. doi:10.1038/nn.4064.
- Maddock, R., Garrett, A., Buonocore, M. (1999). Functions of the retrosplenial cortex in man: emotion, memory, and the “default state” of the brain. *Journal of Cognitive Neuroscience*. 20. Retrieved from <Go to ISI>://WOS:000082700000053.

- Maniscalco, B., Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–30. doi:10.1016/j.concog.2011.09.021.
- Marangoni, C., Garcia, S., Ickes, W., Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology*, *68*(5), 854–69. doi:10.1037//0022-3514.68.5.854.
- McCurdy, L.Y., Maniscalco, B., Metcalfe, J., Liu, K.Y., de Lange, F. P., Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *Journal of Neuroscience*, *33*(5), 1897–906. doi:10.1523/Jneurosci.1890-12.2013.
- Meaux, E., Roux, S., Batty, M. (2013). Early visual ERPs are influenced by individual emotional skills. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nst084.
- Meaux, E., Vuilleumier, P. (2016). Facing mixed emotions: analytic and holistic perception of facial emotion expressions engages separate brain networks. *NeuroImage*, *141*, 154–73.
- Mehrabian, A. (1996). Manual for the Balanced Emotional Empathy Scale (BEES). (Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA, USA 93940).
- Morales, J., Lau, H., Fleming, S.M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *The Journal of Neuroscience*, *38*(14), 3534–46. doi:10.1523/jneurosci.2360-17.2018/jneurosci.2360-17.2018.
- Moran, R., Teodorescu, A.R., Usher, M. (2015). Post choice information integration as a causal determinant of confidence: novel data and a computational account. *Cognitive Psychology*, *78*, 99–147. doi:10.1016/j.cogpsych.2015.01.002.
- Moritz, S., Glascher, J., Sommer, T., Buchel, C., Braus, D.F. (2006). Neural correlates of memory confidence. *NeuroImage*, *33*(4), 1188–93. doi:10.1016/j.neuroimage.2006.08.003.
- N'Diaye, K., Sander, D., Vuilleumier, P. (2009). Self-relevance processing in the human amygdala: gaze direction, facial expression, and emotion intensity. *Emotion*, *9*(6), 798–806. doi:10.1037/a0017845.
- Navajas, J., Hindocha, C., Foda, H., Keramati, M., Latham, P.E., Bahrami, B. (2017). The idiosyncratic nature of confidence. *Nature Human Behaviour*, *1*(11), 810–8. doi:10.1038/s41562-017-0215-1.
- O'Brien, M.L. (1987). Examining the dimensionality of pathological narcissism: factor analysis and construct validity of the O'Brien Multiphasic Narcissism Inventory. *Psychological Reports*, *61*(2), 499–510. doi:10.2466/pr0.1987.61.2.499.
- Pannu, J.K., Kaszniak, A.W. (2005). Metamemory experiments in neurological populations: a review. *Neuropsychology Review*, *15*(3), 105–30. doi:10.1007/s11065-005-7091-6.
- Pleskac, T.J., Busemeyer, J.R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901. doi:10.1037/a0019737.
- Pouget, A., Drugowitsch, J., Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–74. doi:10.1038/nn.4240.
- Pulford, B.D. (1996). *Overconfidence in human judgement*. University of Leicester.
- Ratcliff, R., Starns, J.J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. *Psychological Review*, *120*(3), 697.
- Realo, A., Allik, J., Nolvak, A., et al. (2003). Mind-reading ability: beliefs and performance. *Journal of Research in Personality*, *37*(5), 420–445. doi:10.1016/S0092-6566(03)00021-7.
- Schmitz, T.W., Johnson, S.C. (2007). Relevance to self: a brief review and framework of neural systems underlying appraisal. *Neuroscience and Biobehavioral Reviews*, *31*(4), 585–96. doi:10.1016/j.neubiorev.2006.12.003.
- Setsompop, K., Gagoski, B.A., Polimeni, J.R., Witzel, T., Wedeen, V.J., Wald, L.L. (2012). Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic Resonance in Medicine*, *67*(5), 1210–24. doi:10.1002/mrm.23097.
- Shadlen, M.N., Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927–39. doi:10.1016/j.neuron.2016.04.036.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., Frith, C.D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Sciences*, *18*(4), 186–93. doi:10.1016/j.tics.2014.01.006.
- Sinanaj, I., Cojan, Y., Vuilleumier, P. (2015). Inter-individual variability in metacognitive ability for visuomotor performance and underlying brain structures. *Consciousness and Cognition*, *36*, 327–37. doi:10.1016/j.concog.2015.07.012.
- Smith, M.L., Cottrell, G.W., Gosselin, F., Schyns, P.G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*(3), 184–9. doi:10.1111/j.0956-7976.2005.00801.x.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., Jacobs, G.A. (1983). *Manual for the State-trait Anxiety Inventory*, Palo Alto, CA: Consulting Psychologists Press.
- Valk, S.L., Bernhardt, B.C., Bockler, A., Kanske, P., Singer, T. (2016). Substrates of metacognition on perception and metacognition on higher-order cognition relate to different subsystems of the mentalizing network. *Human Brain Mapping*, *37*(10), 3388–99. doi:10.1002/hbm.23247.
- Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., Valero-Cabre, A. (2014). Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Frontiers in Integrative Neuroscience*, *8*, 66. doi:10.3389/fnint.2014.00066.
- Vossel, S., Geng, J.J., Fink, G.R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist*, *20*(2), 150–59. doi:10.1177/1073858413494269.
- Ward, A.M., Schultz, A.P., Huijbers, W., Van Dijk, K.R., Hedden, T., Sperling, R.A. (2014). The parahippocampal gyrus links the default-mode cortical network with the medial temporal lobe memory system. *Human Brain Mapping*, *35*(3), 1061–73. doi:10.1002/hbm.22234.
- Whitfield-Gabrieli, S., Moran, J.M., Nieto-Castanon, A., Triantafyllou, C., Saxe, R., Gabrieli, J.D.E. (2011). Associations and dissociations between default and self-reference networks in the human brain. *NeuroImage*, *55*(1), 225–32. doi:10.1016/j.neuroimage.2010.11.048.
- Wimmer, G.E., Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*, *338*(6104), 270–73. doi:10.1126/science.1223252.
- Yeung, N., Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–21. doi:10.1098/rstb.2011.0416.
- Zylberberg, A., Roelfsema, P.R., Sigman, M. (2014). Variance misperception explains illusions of confidence in simple perceptual decisions. *Consciousness and Cognition*, *27C*, 246–53. doi:10.1016/j.concog.2014.05.012.