Influence of transient emotional episodes on affective and cognitive theory of mind

Emilie Qiao-Tasserit,[1,2,3] Corrado Corradi-Dell'Acqua,[2,4,5] and Patrik Vuilleumier [1,2,3]

[1]Laboratory of Behavioural Neurology and Imaging of Cognition, Department of Neuroscience, University Medical Center, University of Geneva, Geneva CH-1206, Switzerland
[2]Geneva Neuroscience Center, University of Geneva, Geneva CH-1206, Switzerland
[3]Swiss Center for Affective Sciences, University of Geneva, Geneva CH-1209, Switzerland
[4]Theory of Pain Laboratory, Department of Psychology, Faculty of Psychology and Educational Sciences (FPSE), University of Geneva, Geneva CH-1211, Switzerland
[5]Center for Mind/Brain Sciences (CIMeC), University of Trento, Rovereto IT-38068, Italy
Correspondence should be addressed to Patrik Vuilleumier, Department of Neuroscience, University of Geneva – Campus Biotech, Ch. des Mines 9, Geneva CH-1209, Switzerland. E-mail: patrik.vuilleumier@unige.ch

Abstract

Our emotions may influence how we interact with others. Previous studies have shown an important role of emotion induction in generating empathic reactions towards others' affect. However, it remains unclear whether (and to which extent) our own emotions can influence the ability to infer people's mental states, a process associated with Theory of Mind (ToM) and implicated in the representation of both cognitive (e.g. beliefs and intentions) and affective conditions. We engaged 59 participants in two emotion-induction experiments where they saw joyful, neutral and fearful clips. Subsequently, they were asked to infer other individuals' joy, fear (affective ToM) or beliefs (cognitive ToM) from verbal scenarios. Using functional magnetic resonance imaging, we found that brain activity in the superior temporal gyrus, precuneus and sensorimotor cortices were modulated by the preceding emotional induction, with lower response when the to-be-inferred emotion was incongruent with the one induced in the observer (affective ToM). Instead, we found no effect of emotion induction on the appraisal of people's beliefs (cognitive ToM). These findings are consistent with embodied accounts of affective ToM, whereby our own emotions alter the engagement of key brain regions for social cognition, depending on the compatibility between one's own and others' affect.

Keywords: emotion induction; cognitive theory of mind; affective theory of mind

Introduction

As social creatures, we are able to understand what others are thinking and feeling based on their behavior and context, an ability often referred to as Theory of Mind (ToM). But how our inferential abilities are affected by our own emotional state? Mixed findings have emerged from an heterogeneous literature reporting either no effects of induced emotions on ToM (Holmberg, 2018), decreased performance under positive emotions (Converse *et al.*, 2008) or decreased performance during anxiety but no other negative states (Todd *et al.*, 2015). Unfortunately, these studies vary extensively in terms of methods to induce emotions (clips, music, autobiographical memory), tools to assess ToM (vignettes, visual/conceptual perspective taking, etc.) and nature of the to-be-inferred state (belief, emotion, etc.). Here we begin clarifying the effects of emotions on the brain substrates of ToM by systematically manipulating the affective state inferred in others, under a carefully controlled and validated experimental paradigm for emotion induction and mental state attribution.

Neuroimaging research has unveiled a broad network involved when inferring others' beliefs/thoughts (cognitive ToM), comprising the temporo-parietal junction (TPJ), middle/superior temporal gyrus (STG/MTG), precuneus, lateral and medial prefrontal cortices (PFC) (Van Overwalle and Baetens, 2009; Bzdok *et al.*, 2012; Van Veluw and Chance, 2014; Krall *et al.*, 2015; Molenberghs *et al.*, 2016; Schurz *et al.*, 2017). Importantly, parts of this network also activate when inferring others' emotions (affective ToM), with additional involvement of the temporal pole, amygdala, insula and parts of medial PFC that are not modulated by cognitive ToM (Hynes *et al.*, 2006; Völlm *et al.*, 2006; Sebastian *et al.*, 2012; Corradi-Dell'Acqua *et al.*, 2014; Schlaffke *et al.*, 2015). A differential role of insula/temporal pole and PFC in cognitive and affective ToM has been confirmed by employing neurostimulation (Kalbe *et al.*, 2010) or testing brain-damaged patients (Shamay-Tsoory *et al.*, 2006, 2010; Shamay-Tsoory and Aharon-Peretz, 2007; Corradi-Dell'Acqua *et al.*, 2020). However, a comprehensive data-driven study identified three predominant ToM networks, one cognitive

characterized by TPJ, MTG, PC and PFC; one affective character-ized by temporal pole, amygdala, insula, inferior frontal gyrus and middle cingulate cortex; and one made of partial overlaps in both affective and cognitive nodes (Schurz *et al.*, 2020). To our knowl-edge, no study so far examined whether emotions in the viewer may differentially influence the responsiveness of these networks.

One hypothesis relies on the *embodied* accounts for *empathic sharing*, positing that we might understand emotions of others by mapping them on our own somatic and neural representa-tions (Keysers and Gazzola, 2009; De Waal and Preston, 2017; Ross and Atkinson, 2020). For instance, people in a given emotional state are more sensitive to congruent emotion features in facial expressions (Mobbs *et al.*, 2006; Calbi *et al.*, 2017; Qiao-Tasserit *et al.*, 2017). Likewise, people receiving a pleasant/unpleasant tac-tile stimulation judge incongruent stimulations in others as more neutral (Silani *et al.*, 2013; Riva *et al.*, 2016). Importantly, however, although empathic sharing is usually considered independent from (or even antithetic to) ToM (Dvash and Shamay-Tsoory, 2014; Kanske *et al.*, 2016), advanced studies demonstrated that brain network subserving these two abilities interacted dynamically during naturalistic tasks (Raz *et al.*, 2014). Hence, it is reason-able to assume that one's emotions might influence affective (but not cognitive) ToM abilities consistently with the predictions of embodied accounts, by improving our proficiency in inferring others' state which are consistent with one's own.

Another plausible hypothesis derives from the *Broaden-and-build* theory which posits that positive emotions might broaden attention towards other people and therefore towards their feel-ings and beliefs, while negative emotions may narrow down attention to more personally relevant information and factual details (Fredrickson, 1998; Fredrickson *et al.*, 2004). This frame-work predicts that positive emotions might improve proficiency in any form of ToM, regardless of the state being appraised in others, while negative emotions might decrease it. Indeed, some studies reported that positive emotions increase prosocial behav-iors (Dunn *et al.*, 2008; Aknin *et al.*, 2013a, 2013b) and com-passion (Singer and Klimecki, 2014), while reducing egocentric perspective-taking (Todd *et al.*, 2015). In contrast, negative emo-tions can suppress sensitivity to others' affect, such as pain (Li *et al.*, 2017; Qiao-Tasserit *et al.*, 2018).

The current study directly tested these two opposing predic-tions using a novel paradigm. Individuals were induced with joyful, neutral or fearful emotions through a well-established movie-based procedure previously validated both at the behav-ioral and neural level (Eryilmaz *et al.*, 2011; Pichon *et al.*, 2015; Qiao-Tasserit *et al.*, 2017, 2018). Subsequently, participants per-formed a cognitive and affective ToM task adapted from previous studies (Corradi-Dell'Acqua *et al.*, 2014, 2020), where they read short narratives and then judged the cognitive/affective state of the story's protagonist. The stories described a protagonist in a joyful or fearful situation (affective ToM), in a false-belief situation (cognitive ToM), or as control, a physical object without human protagonist. Across two experiments, we collected behavioral, psychophysiological and neural responses through fMRI (func-tional magnetic resonance imaging) to determine whether emo-tion induction modulated ToM performance and corresponding brain activity. We could thus test the predictions of the *embodied* account (i.e. one's own joy and fear should enhance the pro-cessing of same states in others) and compare them with those of the *Broaden-and-Build* theory (i.e. one's joy enhances sensitiv-ity towards others' states in general, while fear may restrict ToM abilities). The key disambiguating condition concerns changes in the response to fear in others, which *embodied* accounts suggest should be improved by one's fear relative to joy, whereas the Broaden and Build theory suggests the opposite.

Materials and methods
Participants
Twenty-six participants took part in Experiment 1 (15 females, mean age: 27.23, std: 7.79, range: 20–48) and 33 in Experiment 2 (17 females, mean age: 25.33, std: 5.31, range: 18–38). Experi-ment 1 was a preliminary behavioral study testing the paradigm's feasibility. Experiment 2 was the main study and used the same task now in the MRI scanner, combined with psychophysiologi-cal recordings (see Supplementary Appendices). In both experi-ments, participants had no history of psychiatric or neurological disease and gave written informed consent. We conducted this study, approved by the local ethics committee, according to the Declaration of Helsinki.

Stimuli
Movie-Clips
Emotional states were induced by movie-clips: six joyful, six neu-tral and six fearful. This procedure previously demonstrated reli-able emotion induction effects at behavioral (Qiao-Tasserit et al., 2017) and neural levels (Gaviria et al., 2021; Pichon et al., 2015; Qiao-Tasserit et al., 2018; full database description in Eryilmaz et al., 2011). We labelled these three emotion induction conditions based on movie-clips as 'Self-emotion' conditions, in contrast with the ToM manipulation based on stories labelled as 'Others' emo-tion' conditions, where different emotional/cognitive states are inferred in others (see further).

Stories
Participants read 72 short stories in French divided into four cat-egories (Others' Joy, Others' Fear, Others' Belief, Photo) with 18 scenarios each, adapted from previous studies (Saxe and Kan-wisher, 2003; Hynes et al., 2006; Saxe and Powell, 2006; Corradi-Dell'Acqua et al., 2014, 2020). Others' Joy, Others' Fear and Others' Belief stories, describing a protagonist in various situations, were followed by questions probing for his/her emotion (joy/fear) or belief in that situation. We added Photo control stories without protagonist, requiring to infer a change in visual elements (e.g. outdated map). An example of a fearful story would be 'Luc likes to cook for his friends. He prepared a dish that smells very good. When cleaning his kitchen counter, he drops the knife towards his foot. When the knife drops, Luc is… 1) terrified 2) pleased'. Full details about stories are provided in the Supplementary Appendices.

Experimental set-up
Affective and cognitive ToM task. The experiment comprised three scanning sessions, each comprising six mini-blocks (total 18 mini-blocks for an overall of ~45 min). Each mini-block started with a ~1-min emotion induction movie-clip, followed by four stories in random order and then a ToM judgment. Each story was presented for 11 s, followed by an empty screen (2.5–9.5 s, average 5 s), and subsequently by the judgment screen (5 s). At this stage, participants had to select one of two possible story outcomes. The judgment phase was followed by another empty screen (2.5–9.5 s) (Figure 1). To minimize the number of emotional switches, we semi-randomized the order of the blocks within each session, with two movies of the same valence presented consec-utively, and self-joy and self-fear blocks separated by self-neutral blocks.
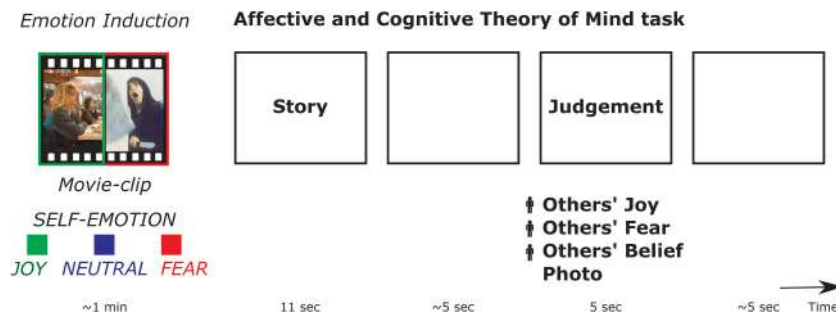
Fig. 1. Experimental set-up. At the beginning of each trial, participants watched a joyful, neutral, or fearful movie-clip (induction of self-emotion). Then they read a short story and gave a forced-choice judgment on whether the protagonist experienced joy, fear (Affective ToM) or a particular belief (Cognitive ToM). We used Photo stories, without human protagonist, as control.

Post-experimental ratings. After scanning, participants watched the beginning of every movie clip and recalled their subjective experience during the main experiment on a visual analog scale. This was achieved through a well-established procedure where we verbally asked participants if they remembered how they felt when they first saw the video-clip, together with scores of plausibility, understandability and how much they felt absorbed during the first watch (Eryilmaz *et al.*, 2011; Pichon *et al.*, 2015; Qiao-Tasserit *et al.*, 2017). For stories with a human protagonist, they also reported the extent to which they reflected upon the protagonist's joy, fear and belief while reading the story (Corradi-Dell'Acqua *et al.*, 2014). See Supplementary Appendices for full details about ratings. 12 (out of 26) participants from Experiment 1, and all 33 participants from Experiment 2, performed the post-experimental procedure, leading to an overall population of $N = 45$ for which these measures were collected.

Apparatus. Stimuli were presented with Matlab and Psychophysics Toolbox (http://psychtoolbox.org/, (Brainard, 1997; Pelli, 1997; Kleiner *et al.*, 2007). In Experiment 1, participants watched them on a PC screen (1024×768 resolution) and responded using the keyboard. In Experiment 2, participants performed the task during fMRI within a 3 T Magnetom TIM Trio scanner (Siemens, Erlangen, Germany) saw stimuli on a LCD screen (CP-SX1350; Hitachi—1024×768 resolution) through a mirror fixed on the MRI head-coil. They answered with key presses on a bimanual button box (HH- 2×4-C; Current Designs). We also recorded electrodermal activity with an MP36R system and SS57L sensors coupled with pre-gelled EL507 electrodes (Biopac Inc, Santa Barbara, CA, USA), and sampled the data at 200 Hz with AcqKnowledge software. Furthermore, we monitored pupil size using an eye-tracker ASL EyeTrac 6 (Applied Science Laboratories, USA) running with a sampling rate of 60 Hz. Finally, we acquired gradient-echo T2*-weighted transverse echo-planar images (EPIs) for blood oxygenation level-dependent (BOLD) contrast, with repetition time 2100 ms, echo time 30 ms, descending acquisition mode, flip angle 80° and in-plane resolution 64×64 voxels (isometric voxel size of 3.2 mm). We also acquired a high-resolution T1-weighted anatomical image. Each volume contained 36 slices of 3.2 mm thickness with a gap of 0.6 mm between slices.

Data processing
Behavioral data. We analyzed behavioral data from each Experiment independently. Using R software version 4.0.0 (https://www.r-project.org/) and R Studio (https://rstudio.com/), for each story category and each movie-clip conditions, we computed individual accuracy scores and median reaction times of correct responses. We fed them into a repeated-measure analysis of variance (ANOVA) with Story Category (*Others' Joy, Others' Fear, Others' Belief* and *Photo*) and Self-Emotion (*Self-Joy, Self-Neutral* and *Self-Fear*) as within-subject factors, followed by *post-hoc* paired *t*-tests. We repeated the same ANOVA on accuracy and reaction time with individual scores from questionnaires of interest included as covariates. For Reaction Time data of Experiment 2, three participants (out of 33) had one missing cell, which was replaced by the average response for the remaining conditions in the same subject. We compared the outcome of this statistical analysis with one obtained when removing these three subjects and found no difference.

We analyzed the post-experimental ratings for movies through a repeated-measure ANOVA testing the effect of Self-Emotion (*Self-Joy, Self-Neutral* and *Self-Fear*), and those for Stories through an ANOVA testing the effect of Story Category. We modeled Story conditions either as a four-level factor (*Others' Joy, Others' Fear, Others' Belief* and *Photo*) for plausibility and understandability judgments, or as a three-level factor (without *Photo*) for joy/fear/belief judgments.

Imaging data

Preprocessing We analyzed MRI images with SPM12 (Wellcome Department of Cognitive Neurology, London, UK). For each participant, we realigned functional images to the first volume of each session. We coregistered these images with the anatomical T1 image, which was then normalized to the standard Montreal Neurological Institute through the unified segmentation approach (Ashburner and Friston, 2005). The normalization deformation field was in turn applied to functional images, which were then resampled to an isotropic voxel size of 2 mm, and spatially smoothed with an isotropic full-width at half-maximum Gaussian kernel of 8 mm.

First-level analyses We fed preprocessed data into a first-level analysis using the General Linear Model framework in SPM12. For each session, we modelled movie, story, and judgement epochs. Movie epochs were modeled in terms of three a boxcar functions (*Self-Joy, Self-Neutral*, and *Self-Fear*) describing blocks with duration of the corresponding movies. Three separate story epochs (following each emotional movie) were modelled as events of 11 s. Reminiscently of Corradi-Dell'Acqua *et al.* (2014), story epochs were combined across the different categories, and associated with four parametric modulation regressors. These included three post-experimental ratings describing the extent to which participants thought about the story protagonist's joy, fear or belief

(*Photo* stories were always coded as 0), and the number of words in each text. To avoid parametrical predictor order biases and to ensure that each effect associated was uniquely interpretable, we modeled story epochs by removing the serial orthogonalization option from SPM default settings. Finally, we modelled 12 judgment epochs (4 Story Categories * 3 Self-Emotions) as events of 5 s. This led to overall 30 predictors (3 movies + 3 stories + 12 story parametric modulations + 12 judgments), each of which was convolved with the canonical Hemodynamic Response Function and associated with its first-order time derivative. We also included six movement parameters as covariates of no interest (*x*, *y*, *z* translations, pitch, roll and yaw rotations). We filtered the low-frequency signal drifts with a cutoff period of 128 s.

Second-level analyses The first-level parameter estimates associated with movie epochs were fed into a second-level flexible factorial design with Self-Emotion (*Self-Joy, Self-Neutral*, and *Self-Fear*) as within-subject factor, and 'Subjects' as random factor, using a random effects analysis. Similarly, the story epochs parameter estimates for each continuous regressor of interest (joy, fear and belief ratings) were modeled through a factorial design with Self-Emotion (*Self-Joy, Self-Neutral* and *Self-Fear*) as within-subject factor. Finally, the parameter estimates associated with judgment events were modeled through a factorial design using one within-subject factor with 12 levels (3 Self-Emotions * 4 Story Categories), and 'Subjects' as random factor. We allowed the within-subjects factor to have unequal variance between their levels, whereas we assumed equal variance for the 'Subjects' factor. We retained significant voxels as those with an extent threshold corresponding to $P < 0.05$ corrected for multiple comparisons (Friston *et al.*, 1993), with an underlying height threshold of $P < 0.001$ (uncorrected).

Results

Movies

In post-experimental testing, participants rated joyful and fearful movies as more arousing and absorbing than neutral movies (Supplementary Appendix A and Supplementary Table S1). As expected, joyful movies elicited more positive feelings than neutral movies and fearful movies more negative feelings. At the physiological level, joyful and fearful movies elicited enhanced electrodermal reactivity during watching, supporting our assumption about effective emotion induction (Supplementary Table S1). At the neural level, emotionally valenced movies (*vs* neutral) engaged a widespread network including bilateral occipito-temporal cortex extending to supramarginal gyrus, precuneus, and dorsal PFC (Eryilmaz *et al.*, 2011), but decreased activity in posterior cingulate cortex (PCC), bilateral TPJ and right frontoparietal operculum (FPO) extending to the posterior insula (PI). Finally, joyful movies selectively activated right supplementary motor cortex and inferior frontal gyrus, extending to STG, relative to both neutral and fearful movies (Supplementary Table S2). No area was selectively activated for fearful movie-clips.
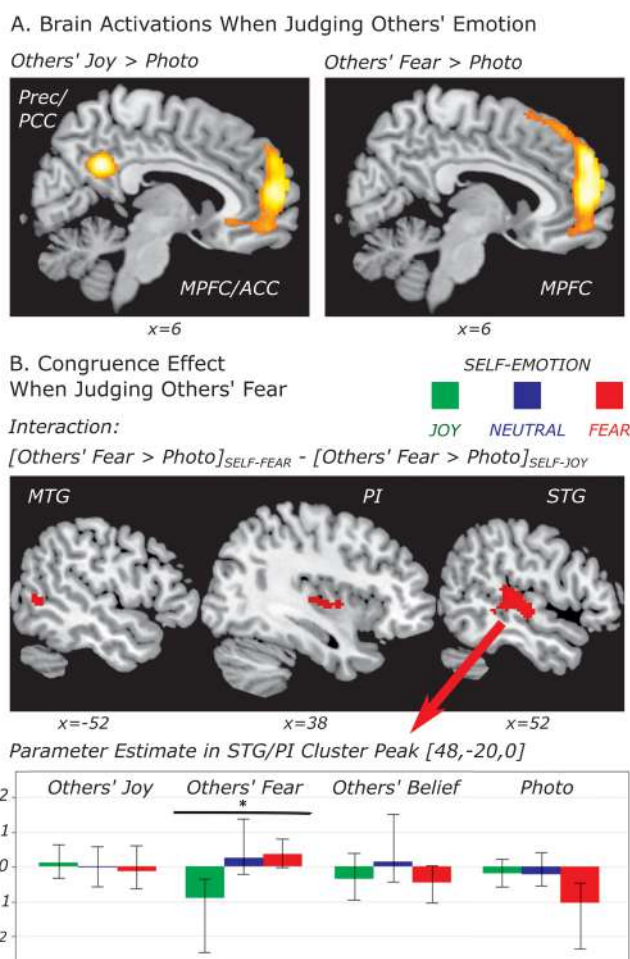
Stories

Others' Joy stories induced the highest ratings of joy and *Others' Fear* stories the highest ratings of fear (Supplementary Table S3). These stories did not differ in beliefs and plausibility ratings. Understandability was equal for all stories in Experiment 1, but slightly higher for *Others' Joy* and slightly lower for *Photo* stories in Experiment 2 (Supplementary Appendix B). At the brain level (Supplementary Table S4), reading about others' joy differentially modulated the right TPJ, while reading about others' fear did not produce any suprathreshold modulation. Instead, scenarios about others' beliefs increased activity in the precuneus and dorsal PFC. These neural activation patterns were not affected by the different Self-emotion conditions.

ToM judgements
Behavioral results

Accuracy was overall high (Experiment 1: 84.48%, std 9.88%; Experiment 2: 84.67%, std 15.49%) (Supplementary Table S3). Two-way repeated-measure ANOVAs (4 Story Categories × 3 Self-Emotions) showed no main effect or interaction associated with *Self-Emotion*, for neither accuracy nor median reaction times ($F \leq 2.10$, n.s.). However, there was a main effect of Story Category for accuracy [Experiment 1: $F_{(3,75)} = 20.97$, $P < 0.001$, Experiment 2: $F_{(3,96)} = 3.80$, $P = 0.01$] and reaction times [Experiment

Fig. 2. Neural response when judging others' joy and others' fear stories. (**A**) The whole-brain map highlights higher activity evoked when judging Others' Joy and Others' Fear relative to control Photo stories. (**B**) Upper panel: activation evoked when judging Others' Fear relative to Photo stories, while being in a congruent self-fear, as compared to an incongruent self-joy emotion. Lower panel: beta parameters of activity extracted from the right STG cluster peak showing a reduced activation when judging Others' Fear, while being in an incongruent self-joy emotion. Green, blue and red bars refer to responses observed after joyful, neutral or fearful movie-clips, respectively, displayed with bootstrap based 95% confidence intervals. 'Asterisk': one-sample *t*-test at $P < 0.05$. Prec: Precuneus. PCC: Posterior Cingulate Cortex. MPFC: Medial Prefrontal Cortex. ACC: Anterior Cingulate Cortex. MTG: Middle Temporal Gyrus. PI: Posterior Insula. STG: Superior Temporal Gyrus.

Table 1. Neural response when judging others' mental content

| | | Coordinates | | | | |
	Side	x	y	z	T	Cluster size
Others' Joy > Photos						
Precuneus/Post. Cingul. Cortex	M	4	−50	30	6.23	645***
Medial Prefrontal Cortex/Ant. Cingul. Cortex	M	8	56	12	5.79	2514***
Others' Fear > Photos						
Medial Prefrontal Cortex/Ant. Cingul. Cortex	M	8	54	10	6.69	3609***
Interaction: [Others' Fear$_{SELF-FEAR}$ > Photo$_{SELF-JOY}$]—[Others' Fear$_{SELF-FEAR}$ > Photo$_{SELF-JOY}$]						
Sup. Temporal Gyrus/Post. Insula	R	48	20	0	4.27	646**
Sup. Occipital/Fusiform Gyrus	L	−20	−78	18	4.27	610**
Middle Temporal Gyrus	L	−50	−60	4	4.33	
Prefrontal Cortex/Caudate	R	48	42	−6	3.85	363*
Congruence: [Others' Joy$_{SELF-JOY}$ + Others' Fear$_{SELF-FEAR}$]—[Others' Fear$_{SELF-JOY}$ + Others' Joy$_{SELF-FEAR}$]						
Sup. Temporal Gyrus/Prec. Gyrus/Post. Insula	R	66	−30	8	4.86	1296***
Sup. Temporal Gyrus/Planum Polare	L	−54	−8	−2	4.74	436*
Frontoparietal operculum	L	−50	−12	8	3.85	
Precuneus	M	6	−52	56	3.86	309*
Calcarine cortex	R	28	−66	8	3.98	635**
Others' Belief > Photos						
Temporoparietal junction	R	54	−58	30	8.25	1081***
Temporoparietal junction	L	−52	−60	30	6.82	1179***
Middle temporal gyrus	R	62	−16	−10	5.52	658**
Middle temporal gyrus	L	−62	−16	−12	5.86	481**
Precuneus/Post. Cingul. Cortex	M	4	−56	32	9.78	1959***
Medial Prefrontal Cortex/Ant. Cingul. Cortex	M	4	52	−10	5.31	2280***

Main activity evoked when judging Others' Joy, Fear and Belief, relative to control Photo stories (without human protagonist). Interaction: brain activation evoked when judging Others' Fear relative to Photo stories, while being oneself in a congruent fearful induced-emotion as compared to an incongruent joyful one. Congruence effect: suppressed brain activity when judging others' emotions (joy/fear), while being oneself in an incongruent induced-emotion (fear/joy) as compared to a congruent one (joy/fear). All clusters survive correction for multiple comparisons at the cluster level (with an underlying height threshold corresponding to P < 0.001, uncorrected). Coordinates in standard MNI space refer to maximally activated foci: x = distance (mm) to the right (+) or the left (−) of the midsagittal line; y = distance anterior (+) or posterior (−) to the vertical plane through the anterior commissure (AC); z = distance above (+) or below (−) the inter-commissural (AC–PC) line. L and R refer to the left and right hemisphere, respectively. M refers to medial activations. Ant.: Anterior. Post.: Posterior. Cingul.: Cingulate. Sup.: Superior. Family-wise significance corrected for multiple comparisons at the cluster level is noted *P < 0.05, ** P < 0.01 and *** P < 0.001.

1: $F_{(3,75)} = 11.43$, $P < 0.001$, Experiment 2: $F_{(3,96)} = 11.82$, $P < 0.001$]. Post-hoc pairedsample *t*-tests revealed that answers for *Others' Joy* stories were equally (Experiment 1) or slightly more accurate (Experiment 2) relative to those for *Others' Fear*, and more accurate relative to *Others' Belief* in both experiments. *Others' Fear* answers were also more (Experiment 1) or equally accurate (Experiment 2) relative to *Others' Belief* answers. The least accurate were the *Photo* stories. Likewise, reaction times were the slowest to *Photo* stories, and equally faster to all other story categories. Post-experimental questionnaire scores, when added as ANOVA covariates, did not play a significant role on accuracy or reaction times [$F_{(6258)} \leq 1.39$, n.s.].

Imaging results

When assessing others' emotional states (affective ToM) regardless of the Self-emotion induction (Figure 2, Table 1), judgments of both fear (*Others' Fear—Photo*) and joy (*Others' Joy—Photo*) relative to the *Photo* condition significantly increased activity in the medial PFC (Figure 2A, Table 1). We also found differential activation in the bilateral TPJ under a less stringent threshold (P < 0.001 uncorrected). In addition, judgments about *Others' joy* also activated the precuneus and the PCC, although the contrast *Others' Joy—Others' Fear* did not show suprathreshold effects.

We then looked at how the judgment about others' affect was influenced by the induced emotional state. The interaction term testing for higher activity associated with fear judgements following congruent (fearful) as opposed to incongruent (joyful) movies ([*Others' Fear > Photo*]$_{SELF-FEAR}$—[*Others' Fear > Photo*]$_{SELF-JOY}$), revealed bilateral increases in superior temporal gyrus, extending to right

PI, as well as in caudate nucleus and occipital cortex. As seen in Figure 2B, these regions showed lower activity during fear-related ToM judgements following incongruent (self-joy) movies compared to congruent (fearful) movies [$t_{(32)} = 2.53$, $P = 0.017$]. When testing for a similar congruency effect of joy-related judgments ([*Others' Joy > Photo*]$_{SELF-JOY}$—[*Others' Joy > Photo*]$_{SELF-FEAR}$), no suprathreshold difference was found.

Finally, we formally tested for a global embodied effect, through a specific two-way interaction contrast ([*Others' Joy$_{SELF-JOY}$ + Others' Fear$_{SELF-FEAR}$*] > [*Others' Joy$_{SELF-FEAR}$ + Others' Fear$_{SELF-JOY}$*]). This again revealed significant modulations in the bilateral STG, extending to FPO bilaterally and to right PI, as well as in the precuneus and medial occipital cortex (Figure 3). In all cases, judgments of others' fear elicited lower activations after exposure to incongruent joyful movies, relative to congruent fearful movies [$t_{(32)} \geq 2.02$, $P \leq 0.05$; marginally significant for the precuneus, $t_{(32)} = 1.79$, $P = 0.08$]; whereas conversely, judgments of others' joy produced lower activity after incongruent fearful movies, relative to congruent joyful or neutral movies [$t_{(32)} \geq 2.35$, $P \leq 0.03$].

Interestingly, in several of these regions, activity was reduced by the incongruent self-emotional conditions compared to the self-neutral one: specifically, the comparison *Others' Fear$_{SELF-NEUTRAL}$ > Others' Fear$_{SELF-JOY}$* also activated right STS [$t_{(32)} = 2.15$, $P = 0.04$] and precuneus [$t_{(32)} = 2.91$, $P = 0.006$], and *Others' Joy$_{SELF-NEUTRAL}$ > Others' Joy$_{SELF-FEAR}$* activated precuneus [$t_{(32)} = 2.10$, $P = 0.04$]. Hence, the interaction between emotions in self and others seemed better explained in terms of incongruence-related deactivation, rather than congruence-related enhancement. Instead, when testing for increased activity associated with
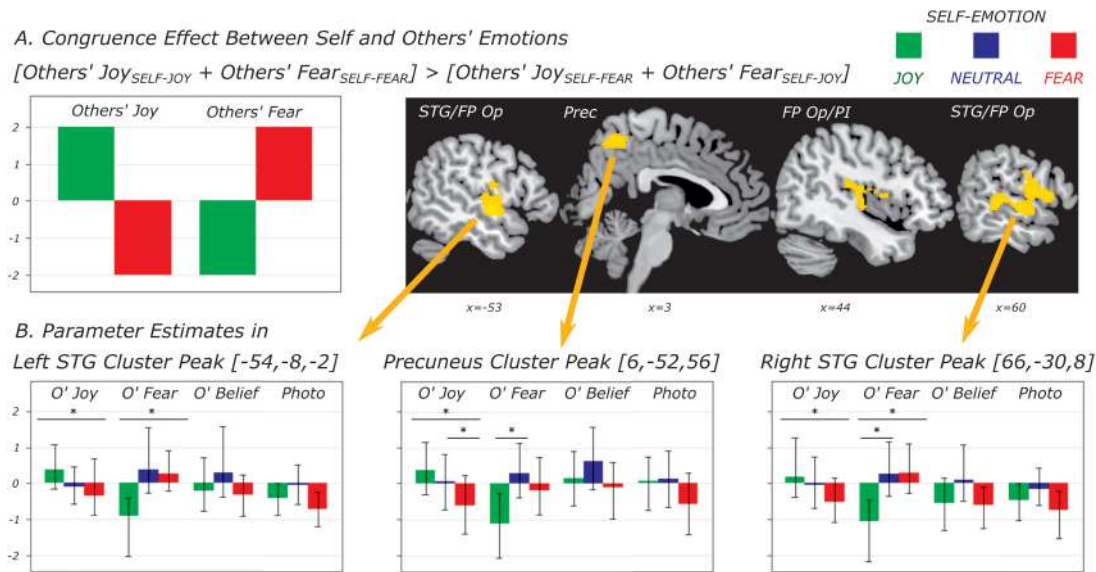
Fig. 3. Neural response to the congruence effect between self-emotion and others' emotions. (**A**) Left panel: Idealized parameter estimates of a congruence effect predicted by the embodiment hypothesis, corresponding to suppressed activity when judging others' emotions (joy/fear), while being oneself in an incongruent induced-emotion (respectively, fear/joy) as compared to a congruent one (respectively, joy/fear). The colored bars depict responses while being in a joyful (green), fearful (red) or neutral (blue) self-emotion, respectively. Right panel: Whole-brain map of brain regions responding to the incongruence effect. (**B**) Parameters extracted from the maximum cluster peak showing an incongruence effect as predicted in (A). Graph bars are displayed with bootstrap based 95% confidence intervals. 'Asterisk': one-sample *t*-test at $P < 0.05$. STG: Superior Temporal Gyrus. FP Op: Frontoparietal Operculum. Prec: Precuneus. PI: Posterior Insula.

emotion incongruence ($[Others'\ Joy_{SELF\text{-}FEAR} + Others'\ Fear_{SELF\text{-}JOY}] > [Others'\ Joy_{SELF\text{-}JOY} + Others'\ Fear_{SELF\text{-}FEAR}]$) showed no suprathreshold activation.

No brain region showed significant modulations compatible with the *Broaden-and-build* hypothesis, i.e. globally enhanced responses after positive self-emotion but reduced after negative self-emotion: $[Others'\ Joy_{SELF\text{-}JOY} + Others'\ Fear_{SELF\text{-}JOY}] > [Others'\ Joy_{SELF\text{-}FEAR} + Others'\ Fear_{SELF\text{-}FEAR}]$).

Finally, the assessment of protagonists' beliefs (cognitive ToM) against control (*Photo* stories), elicited widespread activity in a network involving the precuneus, bilateral TPJ, MTG and medial PFC (Saxe and Kanwisher, 2003; Saxe and Powell, 2006; Mar, 2011; Corradi-Dell'Acqua *et al.*, 2014, 2020). This pattern was observed independently of the induced self-emotion, without any suprathreshold influence on evoked neural responses to beliefs.

Discussion

We found that changes in one's emotions influenced brain responses during affective, but not cognitive ToM, in ways consistent with predictions of *embodied* accounts. Specifically, neural activity evoked when appraising others' fear was larger when individuals were in a fearful state themselves, as opposed to a joyful state ($[Others'\ Fear > Photo]_{SELF\text{-}FEAR} — [Others'\ Fear > Photo]_{SELF\text{-}JOY}$), with predominant impact on the right STG extending to the PI, as well as on the left MTG and right caudate. This effect was confirmed when testing for global embodied effects ($[Others'\ Joy_{SELF\text{-}JOY} + Others'\ Fear_{SELF\text{-}FEAR}] > [Others'\ Joy_{SELF\text{-}FEAR} + Others'\ Fear_{SELF\text{-}JOY}]$), by implicating STG, PI, FPO and precuneus in the appraisal of either joy or fear while being in a congruent (*vs* incongruent) emotional state. We found no evidence supporting the *Broaden and Build* account. Moreover, the impact of self-emotion was selective on mentalizing about affective states, as we found no effect during cognitive ToM. Albeit limited to neural activity,

our results suggest that when inferring affect in others, self-emotions can alter the degree with which we react to others' state, by modulating the responses in specific parts of ToM networks.

Embodied effects when inferring emotions in others

We observed extensive brain activations associated with judgments of joy or fear in others, extending from the ventral to the dorsal medial PFC (Figure 2A) and, for joy judgments especially, in the precuneus. We also observed activation in TPJ under uncorrected threshold. These data dovetail with previous studies suggesting that affective ToM recruits a widespread network involving TPJ, STG/MTG, precuneus, insula and medial PFC, partly overlapping with that observed for cognitive ToM (Hynes *et al.*, 2006; Shamay-Tsoory *et al.*, 2006, 2009; Völlm *et al.*, 2006; Shamay-Tsoory and Aharon-Peretz, 2007; Sebastian *et al.*, 2012; Corradi-Dell'Acqua *et al.*, 2014, 2020; Schlaffke *et al.*, 2015). Remarkably, our new findings reveal that manipulating self-emotion can influence specific nodes within this network.

Brain regions showing emotion congruency effects (displayed in Figure 3) are traditionally associated with different functions. On the one hand, STG and precuneus have been associated with both cognitive and affective ToM (Van Overwalle and Baetens, 2009; Bzdok *et al.*, 2012; Van Veluw and Chance, 2014; Krall *et al.*, 2015; Molenberghs *et al.*, 2016; Schurz *et al.*, 2017, 2020), possibly mediating core ToM processes involved in representing any kind of mental state in others. STG also activates when perceiving socially relevant information (Adolphs, 2001; Pourtois *et al.*, 2004) as well as processing emotional and non-emotional face features (Schobert *et al.*, 2018). In parallel, the precuneus functions extend from ToM to self-processing and episodic memory (Cavanna and Trimble, 2006; Spreng *et al.*, 2009). It possibly provides a hub for the integration of external and internal information, via its widespread connectivity with higher associative regions (Utevsky *et al.*, 2014). Recently, Sharvit *et al.* (2020) implicated a similar

network comprising precuneus, TPJ, STG and medial PFC when individuals assessed the appropriateness of people's conducts, with the precuneus exhibiting high connectivity with an insular portion sensitive to olfactory disgust. Our data extend these previous findings by showing how this region may bridge representations of self-related affective experiences with evaluation of people's mental states.

On the other hand, FPO and insula are implicated in body representation, somatosensory processing, somatic affect and sensorimotor integration (Craig *et al.*, 2000; Farrell *et al.*, 2005; Tsakiris *et al.*, 2007; Corradi-Dell'Acqua *et al.*, 2009; Salimi-Khorshidi *et al.*, 2009; Mengotti *et al.*, 2012; Kropf *et al.*, 2019; Sharvit *et al.*, 2020). These regions may also contribute to understand others' emotions through their face or voice (Kragel and LaBar, 2016), as transient opercular disruption with TMS impairs the recognition of facial expressions (Adolphs *et al.*, 2000; Pourtois *et al.*, 2004). Like the precuneus, the operculum seems to integrate external and internal information, as it supports attention towards our own states such as fear (Straube and Miltner, 2011) or pain (Orenius *et al.*, 2017). Similarly, the PI also encodes sensorimotor signals at the interface between external (e.g. touch, temperature, pain, and sounds) and internal (e.g. somatovisceral, vestibular) information (Chang *et al.*, 2013). Finally, the ventral precentral/premotor area may provide an interface between others' perception and self-performed actions as it is thought to carry mirror properties, activated both when observing and when performing goal-directed (Chouinard and Paus, 2006; Morin and Grèzes, 2008; Kantak *et al.*, 2012) or defensive actions (Cooke and Graziano, 2004; De Gelder *et al.*, 2004). Its dysfunction has been associated with social deficits including those related to ToM abilities (Nebel *et al.*, 2014; Díez-Cirarda *et al.*, 2015).

All these regions represent plausible neuronal outputs for embodied influences on affective ToM. A wealthy literature has hypothesized that we represent others' emotions through a process of simulation, whereby we re-map the movement/state observed in others in ourselves (Keysers and Gazzola, 2009; De Waal and Preston, 2017; Ross and Atkinson, 2020). Consequently, our effects are consistent with the 'somatic marker hypothesis', which suggests that emotions are expressed through changes in the representation of somatic/interoceptive states, underpinned by neural activity in insular/opercular structures (Damasio *et al.*, 1996; Craig, 2009). Here, this marker might help appraise others' emotion more efficiently when they are consistent with our own current affective state and deactivated when inconsistent.

To our knowledge, although embodied effects have previously been described for emotional facial processing and pain empathy, this is the first study extending such effects to ToM abilities. Seminal accounts have described at least two neural pathways subserving the processing of others' affect. An 'affective' pathway, mediated by regions such as the insula, is grounded in the representation of one's own affect and bodily states. A 'cognitive' pathway, mediated by a network containing medial PFC, TPJ, precuneus and STG, is associated with cognitive ToM, underlying knowledge about others' mental states, such as beliefs, intentions, etc. (Shamay-Tsoory, 2011; Stietz *et al.*, 2019). However, a recent meta-analysis argued that affective ToM recruits a hybrid network, combining aspects of both 'cognitive' and 'affective' pathways (Schurz *et al.*, 2020). Our data supports this claim by showing how appraisal about others' emotions (I) is influenced by one's own affect, in agreement with *embodied* accounts; (ii) encompasses regions from both 'cognitive' (precuneus and STG) and 'affective' (insula, operculum and ventral premotor cortex)

networks. In this view, the brain allows combining these pathways for a more comprehensive representation of others' affect.

Broaden and Build vs Embodiment

We found no support for the *Broaden-and-Build* hypothesis. This account posits that positive emotions broaden one's attention resources (Fredrickson *et al.*, 2004; Rowe *et al.*, 2007), boosting the ability to infer others' feelings, emotions and beliefs, while negative emotions impair it. Unlike here, we previously found support for this hypothesis in a study where negative emotions suppressed brain activity related to empathy for pain (Qiao-Tasserit *et al.*, 2018). Hence, in at least some conditions, the *Broaden-and-Build* might offer a valid model for explaining sensitivity to others' affect or bodily state, and concomitant modulations of brain activity.
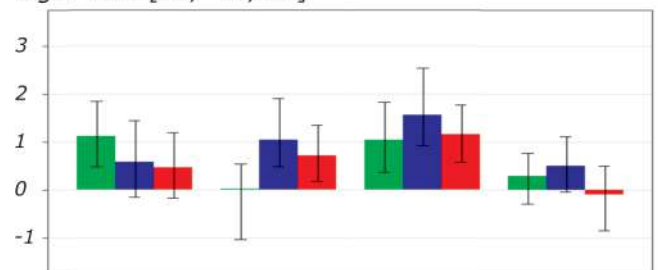
The present study is not equipped to inform why *Embodied* or *Broaden-and-Build* accounts prevail in different circumstances.
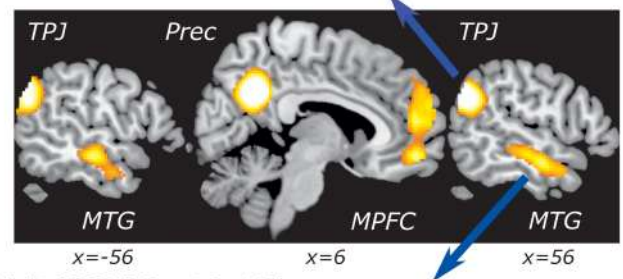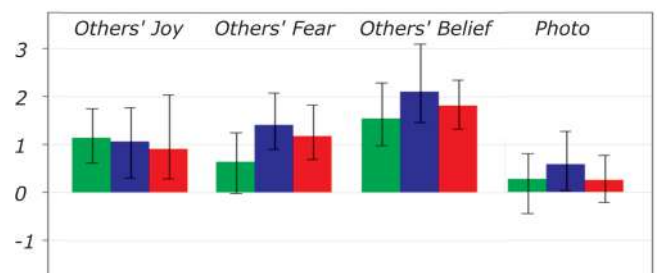
Fig. 4. Neural response when judging others' belief stories. Middle panel: the whole-brain map highlights the main activity evoked when judging Others' Belief relative to control Photo stories (without human protagonist). Upper and lower panels: the parameters extracted from the right TPJ and MTG cluster maximum showed no significant effect of emotion induction. Green, blue and red bars refer to responses provided after a joyful, neutral or fearful movie-clip, respectively, and are displayed with bootstrap based 95% confidence intervals. TPJ: Temporoparietal Junction. Prec: Precuneus. MTG: Middle Temporal Gyrus. MPFC: Medial Prefrontal Cortex.

However, previous studies reported an embodied effect when rating others' pain (Rütgen *et al.*, 2015; Antico *et al.*, 2019) or recognizing facial expressions (Mobbs *et al.*, 2006; Calbi *et al.*, 2017; Qiao-Tasserit *et al.*, 2017). In all these cases, the affective state induced in the observer was the same as the one to be inferred in others (e.g. participants were induced with joy or fear and had to judge joy or fear in others; Qiao-Tasserit *et al.*, 2017). Instead, the *Broaden-and-Build* theory was supported by conditions where the state induced in our participants was qualitatively different from that they observed in others (participants were induced with joy or fear, and had to judge physical pain in others; Qiao-Tasserit *et al.*, 2018). It is therefore possible that an embodied representation to appraise another person's affect requires a close match between one's own and the others' states (e.g. self-fear, others' fear), and/or relies on an indirect inference of subjective experience, in the absence of which (e.g. self-fear, others' sadness) the brain may adopt or favor alternative processing strategies.

Limitations of the study and conclusion

Our study is not without limitations. First, emotional influences on ToM brain mechanisms appeared too subtle to impact explicit behavior measures, potentially reflecting a ceiling effect as suggested by high accuracy overall. Second, our design could not tell whether the embodied effect observed here taps onto a representation of specific emotion states or a broader representation of valence. Future studies involving the assessment of others' affective states carefully matched for this core dimension should shed light on this matter (e.g. does fear induction influence equally the assessment of fear or anger?). Third, post-scanning ratings revealed occasional, though minor, negative appraisals associated with happy stories, whereas no positive appraisals were associated with negative stories (Supplementary Table S3). Luckily, this imbalance in experimental materials did not seem to impact our results, as embodied effects were observed for both joy and fear TOM judgments (Figure 4).

Keeping these limitations aside, we provide conclusive evidence that our own emotions influence how our brain activates to assess those of other people. We found that when inferring emotions in others that are incongruent with our own current affective state, brain activity is reduced in a widespread network comprehending precuneus, STG, insula and parietal operculum. These results help disentangle between opposing theoretical accounts, by demonstrating how our own emotional state can affect the way we embody and simulate the emotions of other people. To conclude, we speculate that being more aware of the way our environment and feelings shape how we respond to others might foster social processes, contributing to more peaceful and gratifying relationships.

Supplementary data

Supplementary data is available at SCAN online.

Data Availability

The datasets analyzed for this study will be made available from the authors on request.

Funding

Conflict of interest

The authors declared that they had no conflict of interest with respect to their authorship or the publication of this article.

Acknowledgements

References

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, **11**(2), 231–9.

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., Damasio, A.R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *The Journal of Neuroscience*, **20**(7), 2683–90.

Aknin, L.B., Barrington-Leigh, C.P., Dunn, E.W., *et al.* (2013a). Prosocial spending and well-being: cross-cultural evidence for a psychological universal. *Journal of Personality and Social Psychology*, **104**(4), 635–52.

Aknin, L.B., Dunn, E.W., Whillans, A.V., Grant, A.M., Norton, M.I. (2013b). Making a difference matters: impact unlocks the emotional benefits of prosocial spending. *Journal of Economic Behavior & Organization*, **88**, 90–5.

Antico, L., Cataldo, E., Corradi-Dell'Acqua, C. (2019). Does my pain affect your disgust? Cross-modal influence of first-hand aversive experiences in the appraisal of others' facial expressions. *European Journal of Pain*, **23**(7), 1283–96.

Ashburner, J., Friston, K.J. (2005). Unified segmentation. *NeuroImage*, **26**(3), 839–51.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–6.

Bzdok, D., Schilbach, L., Vogeley, K., *et al.* (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure & Function*, **217**(4), 783–96.

Calbi, M., Heimann, K., Barratt, D., Siri, F., Umiltà, M.A., Gallese, V. (2017). How context influences our perception of emotional faces: a behavioral study on the Kuleshov effect. *Frontiers in Psychology*, **8**, 1684.

Cavanna, A.E., Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, **129**(3), 564–83.

Chang, L.J., Yarkoni, T., Khaw, M.W., Sanfey, A.G. (2013). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cerebral Cortex*, **23**(3), 739–49.

Chouinard, P.A., Paus, T. (2006). The primary motor and premotor areas of the human cerebral cortex. *The Neuroscientist*, **12**(2), 143–52.

Converse, B.A., Lin, S., Keysar, B., Epley, N. (2008). In the mood to get over yourself: mood affects theory-of-mind use. *Emotion*, **8**(5): 725–30.

Cooke, D.F., Graziano, M.S.A. (2004). Sensorimotor integration in the precentral gyrus: polysensory neurons and defensive movements. *Journal of Neurophysiology*, **91**(4), 1648–60.

Corradi-Dell'Acqua, C., Hofstetter, C., Vuilleumier, P. (2014). Cognitive and affective theory of mind share the same local patterns of activity in posterior temporal but not medial prefrontal cortex. *Social Cognitive & Affective Neuroscience*, **9**(8), 1175–84.

Corradi-Dell'Acqua, C., Ronchi, R., Thomasson, M., Bernati, T., Saj, A., Vuilleumier, P. (2020). Deficits in cognitive and affective theory of mind relate to dissociated lesion patterns in prefrontal and insular cortex. *Cortex*, **128**, 218–33.

Corradi-Dell'Acqua, C., Tomasino, B., Fink, G.R. (2009). What is the position of an arm relative to the body? Neural correlates of body schema and body structural description. *Journal of Neuroscience*, **29**(13), 4162–71.

Craig, A.D.B. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews, Neuroscience*, **10**, 59–70.

Craig, A.D.B., Chen, K., Bandy, D., Reiman, E.M. (2000). Thermosensory activation of insular cortex. *Nature Neuroscience*, **3**(2), 184–90.

Damasio, A.R., Everitt, B.J., Bishop, D., Damasio, A.R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions: Biological Sciences*, **351**(1346), 1413–20.

De Gelder, B., Snyder, J., Greve, D., Gerard, G., Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(47), 16701–6.

De Waal, F.B.M., Preston, S.D. (2017). Mammalian empathy: behavioural manifestations and neural basis. *Nature Reviews, Neuroscience*, **18**(8), 498–509.

Díez-Cirarda, M., Ojeda, N., Peña, J., *et al.* (2015). Neuroanatomical correlates of theory of mind deficit in Parkinson's disease: a multimodal imaging study. *PLoS ONE*, **10**(11), 1–16.

Dunn, E.W., Aknin, L.B., Norton, M.I. (2008). Spending money on others promotes happiness. *Science*, **319**(5870), 1687–8.

Dvash, J., Shamay-Tsoory, S.G. (2014). Theory of mind and empathy as multidimensional constructs: neurological foundations. *Topics in Language Disorders*, **34**(4), 282–95.

Eryilmaz, H., Van De Ville, D., Schwartz, S., Vuilleumier, P. (2011). Impact of transient emotions on functional connectivity during subsequent resting state: a wavelet correlation approach. *NeuroImage*, **54**(3), 2481–91.

Farrell, M.J., Laird, A.R., Egan, G.F. (2005). Brain activity associated with painfully hot stimuli applied to the upper limb: a meta-analysis. *Human Brain Mapping*, **25**(1), 129–39.

Fredrickson, B.L. (1998). What good are positive emotions? *Review of General Psychology*, **2**(3), 300–19.

Fredrickson, B.L., Huppert, F.A., Baylis, N., Keverne, B. (2004). The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, **359**(1449), 1367–78.

Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., Evans, A.C. (1993). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, **1**(3), 210–20.

Gaviria, J., Rey, G., Bolton, T., Ville, D.V.D., Vuilleumier, P. (2021). Dynamic functional brain networks underlying the temporal inertia of negative emotions. *NeuroImage*, **240**, 118377.

Holmberg, P. (2018). The effect of emotional states on theory of mind. Masters' Thesis, August, 1–87.

Hynes, C.A., Baird, A.A., Grafton, S.T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia*, **44**(3), 374–83.

Kalbe, E., Schlegel, M., Sack, A.T., *et al.* (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex*, **46**(6), 769–80.

Kanske, P., Böckler, A., Trautwein, F.-M., Parianen Lesemann, F.H., Singer, T. (2016). Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive & Affective Neuroscience*, **11**(9), 1383–92.

Kantak, S.S., Stinear, J.W., Buch, E.R., Cohen, L.G. (2012). Rewiring the brain: potential role of the premotor cortex in motor control, learning, and recovery of function following brain injury. *Neurorehabilitation & Neural Repair*, **26**(3), 282–92.

Keysers, C., Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, **19**(6), 666–71.

Kleiner, M., Brainard, D.H., Pelli, D.G., Ingling, A., Murray, R., Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, **36**(14), 1–16.

Kragel, P.A., LaBar, K.S. (2016). Decoding the nature of emotion in the brain. *Trends in Cognitive Sciences*, **20**(6), 444–55.

Krall, S.C., Rottschy, C., Oberwelland, E., *et al.* (2015). The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Structure & Function*, **220**(2), 587–604.

Kropf, E., Syan, S.K., Minuzzi, L., Frey, B.N. (2019). From anatomy to function: the role of the somatosensory cortex in emotional regulation. *Brazilian Journal of Psychiatry*, **41**(3), 261–9.

Li, X., Meng, X., Li, H., Yang, J., Yuan, J. (2017). The impact of mood on empathy for pain: evidence from an EEG study. *Psychophysiology*, **54**(9), 1311–22.

Mar, R.A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, **62**(1), 103–34.

Mengotti, P., Corradi-Dell'Acqua, C., Rumiati, R.I. (2012). Imitation components in the human brain: An fMRI study. *NeuroImage*, **59**(2), 1622–30.

Mobbs, D., Weiskopf, N., Lau, H.C., Featherstone, E., Dolan, R.J., Frith, C.D. (2006). The Kuleshov effect: the influence of contextual framing on emotional attributions. *Social Cognitive & Affective Neuroscience*, **1**(2), 95–106.

Molenberghs, P., Johnson, H., Henry, J.D., Mattingley, J.B. (2016). Understanding the minds of others: a neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, **65**, 276–91.

Morin, O., Grèzes, J. (2008). What is "mirror" in the premotor cortex? A review. *Neurophysiologie Clinique*, **38**(3), 189–95.

Nebel, M.B., Eloyan, A., Barber, A.D., Mostofsky, S.H. (2014). Precentral gyrus functional connectivity signatures of autism. *Frontiers in Systems Neuroscience*, **8**, 1–11.

Orenius, T.I., Raij, T.T., Nuortimo, A., Näätänen, P., Lipsanen, J., Karlsson, H. (2017). The interaction of emotion and pain in the insula and secondary somatosensory cortex. *Neuroscience*, **349**, 185–94.

Pelli, D.G. (1997). The video toolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, **10**(4), 437–42.

Pichon, S., Miendlarzewska, E.A., Eryilmaz, H., Vuilleumier, P. (2015). Cumulative activation during positive and negative events and state anxiety predicts subsequent inertia of amygdala reactivity. *Social Cognitive & Affective Neuroscience*, **10**(2), 180–90.

Pourtois, G., Sander, D., Andres, M., *et al.* (2004). Dissociable roles of the human somatosensory and superior temporal cortices for processing social face signals. *European Journal of Neuroscience*, **20**(12), 3507–15.

Qiao-Tasserit, E., Corradi-Dell'Acqua, C., Vuilleumier, P. (2018). The good, the bad, and the suffering. Transient emotional episodes modulate the neural circuits of pain and empathy. *Neuropsychologia*, **116**, 99–116.

Qiao-Tasserit, E., Garcia Quesada, M., Antico, L., Bavelier, D., Vuilleumier, P., Pichon, S. (2017). Transient emotional events and individual affective traits affect emotion recognition in a perceptual decision-making task. *Plos One*, **12**(2), e0171375.

Raz, G., Jacob, Y., Gonen, T., *et al.* (2014). Cry for her or cry with her: context-dependent dissociation of two modes of cinematic empathy reflected in network cohesion dynamics. *Social Cognitive & Affective Neuroscience*, **9**(1), 30–8.

Riva, F., Triscoli, C., Lamm, C., Carnaghi, A., Silani, G. (2016). Emotional egocentricity bias across the life-span. *Frontiers in Aging Neuroscience*, **8**, 1–7.

Ross, P., Atkinson, A.P. (2020). Expanding simulation models of emotional understanding: the case for different modalities, body-state simulation prominence, and developmental trajectories. *Frontiers in Psychology*, **11**, 1–21.

Rowe, G., Hirsh, J.B., Anderson, A.K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(1), 383–8.

Rütgen, M., Seidel, E.-M., Silani, G., *et al.* (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proceedings of the National Academy of Sciences*, **112**(41), E5638–E5646.

Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E. (2009). Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. *NeuroImage*, **45**(3), 810–23.

Saxe, R., Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind. *NeuroImage*, **19**(4), 1835–42.

Saxe, R., Powell, L.J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*, **17**(8), 692–9.

Schlaffke, L., Lissek, S., Lenz, M., *et al.* (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: a neuroimaging study using cartoon picture stories. *Human Brain Mapping*, **36**(1), 29–39.

Schobert, A.K., Corradi-Dell'Acqua, C., Frühholz, S., van der Zwaag, W., Vuilleumier, P. (2018). Functional organization of face processing in the human superior temporal sulcus: a 7T high-resolution fMRI study. *Social Cognitive & Affective Neuroscience*, **13**(1), 102–13.

Schurz, M., Radua, J., Tholen, M.G., *et al.* (2020). Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, **147**, 293–327.

Schurz, M., Tholen, M.G., Perner, J., Mars, R.B., Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: a review using probabilistic atlases from different imaging modalities. *Human Brain Mapping*, **38**(9), 4788–805.

Sebastian, C.L., Fontaine, N.M.G., Bird, G., *et al.* (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive & Affective Neuroscience*, **7**(1), 53–63.

Shamay-Tsoory, S.G. (2011). The neural bases for empathy. *Neuroscientist*, **17**(1), 18–24.

Shamay-Tsoory, S.G., Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia*, **45**(13), 3054–67.

Shamay-Tsoory, S.G., Aharon-Peretz, J., Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, **132**(3), 617–27.

Shamay-Tsoory, S.G., Harari, H., Aharon-Peretz, J., Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, **46**(5), 668–77.

Shamay-Tsoory, S.G., Tibi-Elhanany, Y., Aharon-Peretz, J. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Social Neuroscience*, **1**(3–4), 149–66.

Sharvit, G., Lin, E., Vuilleumier, P., Corradi-Dell'Acqua, C. (2020). Does inappropriate behavior hurt or stink? The interplay between neural representations of somatic experiences and moral decisions. *Science Advances*, **6**(42), eaat4390.

Silani, G., Lamm, C., Ruff, C.C., Singer, T. (2013). Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. *Journal of Neuroscience*, **33**(39), 15466–76.

Singer, T., Klimecki, O.M. (2014). Empathy and compassion. *Current Biology*, **24**(18), R875–R878.

Spreng, R.N., Mar, R.A., Kim, A.S.N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience*, **21**(3), 489–510.

Stietz, J., Jauk, E., Krach, S., Kanske, P. (2019). Dissociating empathy from perspective-taking: evidence from intra- and inter-individual differences research. *Frontiers in Psychiatry*, **10**.

Straube, T., Miltner, W.H.R. (2011). Attention to aversive emotion and specific activation of the right insula and right somatosensory cortex. *NeuroImage*, **54**(3), 2534–8.

Todd, A.R., Forstmann, M., Burgmer, P., Brooks, A.W., Galinsky, A.D. (2015). Anxious and egocentric: how specific emotions influence perspective taking. *Journal of Experimental Psychology: General*, **144**(2), 374–91.

Tsakiris, M., Hesse, M.D., Boy, C., Haggard, P., Fink, G.R. (2007). Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cerebral Cortex*, **17**(10), 2235–44.

Utevsky, A.V., Smith, D.V., Huettel, S.A. (2014). Precuneus is a functional core of the default-mode network. *Journal of Neuroscience*, **34**(3), 932–40.

Van Overwalle, F., Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, **48**(3), 564–84.

Van Veluw, S.J., Chance, S.A. (2014). Differentiating between self and others: an ALE meta-analysis of fMRI studies of self-recognition and theory of mind. *Brain Imaging and Behavior*, **8**, 24–38.

Völlm, B.A., Taylor, A.N.W., Richardson, P., *et al.* (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, **29**(1), 90–8.