

Cry for her or cry with her: context-dependent dissociation of two modes of cinematic empathy reflected in network cohesion dynamics

Gal Raz,^{1,2,3,*} Yael Jacob,^{1,4} Tal Gonen,^{1,5} Yonatan Winetraub,^{1,6} Tamar Flash,⁷ Eyal Soreq,^{1,3,8} and Talma Hendler^{1,2,6}

¹Functional Brain Imaging Unit, Functional Brain Center, Wohl institute for Advanced Imaging, Tel Aviv Sourasky Medical Center, 6 Weizmann Street, Tel Aviv 6423906, Israel, ²Department of Physiology and Pharmacology, Sackler Faculty of Medicine, ³Film and Television Department, ⁴Sagol School of Neuroscience, ⁵School of Psychological Sciences, ⁶Adi Lautman Interdisciplinary Program for Outstanding Students, Tel Aviv University, 30 Haim Levanon street, 6997543, Israel, ⁷Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, 234 Herzl Streer, Rehovot 76100, Israel and ⁸Screen Based Art Department, Bezalel Academy of Arts and Design, Mount Scopus, P.O. Box 24046, Jerusalem 91240, Israel

Two empathy-related processes were recently distinguished neuroscientifically: automatic embodied-simulation (ES) based on visceromotor representation of another's affective state via cingulo-insular circuit, and emotional sharing relying on cognitive 'theory of mind' (ToM) via prefrontal-temporo-parietal circuit. Evidence that these regions are not only activated but also function as networks during empathic experience has yet to be shown. Employing a novel approach by analyzing fMRI fluctuations of network cohesion while viewing films portraying personal loss, this study demonstrates increased connectivity during empathic engagement (probed by behavioral and parasympathetic indices) both within these circuits, and between them and a set of limbic regions. Notably, this effect was context-dependent: when witnessing as a determined-loss presented as a future event, the ToM and ToM-limbic cohesion positively correlated with state- and empathy indices. During the dramatic peak of this condition, the ToM cohesion was positively correlated with the trait-empathy index of personal distress. However, when the loss was presented as a probabilistic real-time occurrence, ToM cohesion negatively correlated with state-empathy index, which positively correlated with ES-limbic cohesion. In this case, it was the ES-limbic cohesion during the emotional peak which was correlated with personal distress scores. The findings indicate a dichotomy between regulated empathy toward determined-loss and vicarious empathy toward a real-time occurrence.

Keywords: embodied simulation; emotion regulation; functional connectivity; network dynamics; theory of mind

INTRODUCTION

Like bodily motion, emotion involves the unfolding coordination of physiological systems. Accordingly, the dynamism of neural connectivity has attracted increasing attention in contemporary affective neuroscience. Prominent theories in the field (e.g. Barrett *et al.*, 2007; Scherer, 2009) maintain that emotional experience emerges out of the ongoing interaction between specialized neural circuits rather than the occasional 'on/off switching' of preset 'affect programs'. According to this view, emotions are dynamically shaped by the convergence and divergence of continuous and fluctuating streams of multi-modal information. This theoretical agenda is congruent with an empirical approach, which is sensitive not only to the intensity of the activity of certain neural structures but also to the dynamic interaction between distinct brain modules.

Nonetheless, both dynamism and connectivity are under-represented in empirical neuroimaging literature on *empathy*. As empathic feelings, like other emotions, are based on the unfolding processing of multi-modal perceptual, social, cognitive and affective cues, they may arise and decay in correspondence with the connectivity level of specific neural circuits. Nevertheless, studies thus far have

largely focused on averaged short-lived and univariate effects of empathy-related stimuli on signal levels. Only few efforts to examine connectivity in fMRI research of empathy were made (Jabbi *et al.*, 2008; Lombardo *et al.*, 2009), and to the best of our knowledge, none of them involved the analysis of the development of functional connectivity in time.

This said, over the past two decades dramatic advances have been made in the psychological research of empathy—a notion, which generally refers to the understanding and sharing of another's emotional state (Eisenberg and Strayer, 1987). Converging evidence suggests that empathy is constructed of at least two separate processes: *embodied simulation* (ES; Gallese, 2005) and *theory of mind* (ToM; Baron-Cohen *et al.*, 1985).

ES refers to an automatic process of vicariously sharing or 'resonating' a bodily state of an observed agent. It is considered a bottom-up process based on the interoceptive representation of another's state. This form of affective mimicry seems to rely on neural circuits involved both in the self-experience of internal body states and the perception of similar states in others. It has been theoretically linked to the concept of the mirror neuron, especially in the context of pain (e.g. Gallese, 2007; Bruneau *et al.*, 2012, but cf. Danziger *et al.*, 2009). A set of regions, encompassing the anterior insula (AI), an adjacent section of the inferior frontal gyrus and the middle anterior cingulate (ACC), were found to be activated both when experiencing a first-hand painful stimulus and when anticipating or observing other's physical pain. A variety of neuroanatomical and functional evidence suggests that these regions together comprise part of a tightly connected network, supporting interoceptive awareness and one's sense of subjectivity (Craig, 2008, 2009). Importantly, mirroring-like effects have been demonstrated in these regions not only in the case of

Received 1 December 2012; Accepted 28 March 2013

Advance Access publication 24 April 2013

The authors thank Yoav Benjamini (Department of Statistics and Operation Research, Tel Aviv University) for helpful statistical consultation, Gadi Gilam for help in scanning, Adi Maron-Kats for computational assistance, and Aliya Sloski for copy editing. The authors also thank Claus Lamm, Danilo Bzdok, and Simon B. Eickhoff for providing meta-analytical data. This study was supported by the Communication with Emotional Body Language (COBOL) European Commission grant 043403, University of Chicago's Arete Initiative and Dan David Scholarship.

Correspondence should be addressed to Gal Raz, Zamenhof 16, Tel-Aviv, 64373, Israel. E-mail: galraz@post-tau.ac.il and to Talma Hendler Weizman 6 Tel Aviv 64239. E-mail: talma@tlvmc.gov.il

physical pain but also in other social aversive contexts (Eisenberger *et al.*, 2003; Wicker *et al.*, 2003; Harrison *et al.*, 2006; Meyer *et al.*, 2012).

On the other hand, ToM relies on *cognitive* rather than interoceptive representations of another's state. These representations allow for a top-down inference of another's mental state by attributing certain beliefs, thoughts, motivations and desires to that person. A distinct set of brain regions has been implicated in this cognitive mode of perspective-sharing, including the ventral and dorsal aspects of the medial prefrontal (MPF) cortex, the superior temporal sulcus (STS) and the temporo-parietal junction (TPJ). The STS and the TPJ have been linked with transient detection and evaluation of information of another's state (Gallagher and Frith, 2003; Van Overwalle and Baetens, 2009), while MPF structures have been associated with longer lasting processing of the self and other related content (Shamay-Tsoory, 2011).

It should be noted that the notion *cognitive* is not taken here as an antonym of the term *affective*. Theoretical accounts of emotion (e.g. Lazarus, 1994; Pessoa, 2008; Russell, 2009) maintain that cognitive processing is an inseparable aspect of emotion, while comprehensive meta-analyses of neuroimaging studies (e.g. Kober *et al.*, 2008; Lindquist *et al.*, 2012) indicate significant involvement of ToM-related regions in emotional experience. Moreover, the MPF-temporo-parietal circuit has been implicated not only in 'cold' reasoning but also in affective empathic reactions based on inferences on other's state even when no significant activation of ES-related regions was evident (e.g. Krämer *et al.*, 2010; Meyer *et al.*, 2012). Thus, the ES–ToM distinction is not equivalent to the affective/non-affective empathy dichotomy. ES- and ToM-related circuits are assumed to have distinctive anatomical connectivity profiles and evolutionary and ontogenetic histories (De Waal, 2008; Decety, 2011), which qualify them as systems specialized in processing different types of information. Although interoception or cognition may often not develop into a full-blown emotional experience, under certain conditions these processes may also drive inter-subjective sharing of emotions as they integrate with relevant input from other perceptual and limbic domains. The relative contributions of each of these systems and their interactions with limbic structures to one's empathic reaction are within the main focus of this study.

Although evidence supporting and elaborating the ES–ToM dissociation accumulates, a recent review of neuroscientific research (Zaki and Ochsner, 2012) highlights typical pitfalls in this literature. In short, it stresses the importance of using realistic experimental manipulations, rich in both perceptual and cognitive cues, to enhance the interpretability of findings and their relevance to real-life empathy. It is also recommended to relate the neural data to corresponding behavioral measures to ensure the participants indeed experienced empathy.

In keeping with the recommendation to employ holistic rather than isolated stimuli when examining empathy-related processes, cinematic stimuli were used in this study. Films are ecological manipulations, as they introduce multi-modal, dynamic and realistic stimuli. They are considered particularly effective in inducing empathy (Smith, 1995). Two excerpts, presenting a fatal separation of a mother from her children, were taken from the films *Stepmom* (1998) and *Sophie's Choice* (1982, referred to here as *Sophie*). Both excerpts were previously reported to effectively induce empathy-related emotions such as sadness (Oatley, 1996; Goldin *et al.*, 2005; Raz *et al.*, 2012), compassion and mercy (Raz *et al.*, 2012).

The dynamic nature of these clips enables the examination of the aforementioned key aspect of emotion highlighted by neuroscientific theories: dynamics of connectivity. This study aims to support the notion that ES- and ToM-related sets of regions actually function as

networks, with inter-regional crosstalk increasing as individuals become empathically engaged.

The empathy-related networks (Figure 1 and Supplementary Table S1) were defined here based on recent meta-analyses of the literature mentioned above (Lamm *et al.*, 2011; Bzdok *et al.*, 2012; see 'Materials and Methods' section for region selection rationale). Since empathy is basically an emotional phenomenon whose evolutionary value relies on its ability to recruit affective-motivational processes (Batson *et al.*, 1981), a third set of subcortical regions implicated in such processes was added to the analysis. This set (i.e. limbic network) includes the periaqueductal gray (PAG), amygdala, ventral striatum, hypothalamus and medial thalamus (Figure 1) that were reported by a third meta-analysis to co-activate across numerous studies of emotion (Kober *et al.*, 2008). Thus, both patterns of connectivity within ES and ToM networks as well as between these networks and limbic regions were examined in this study.

To dynamically probe functional connectivity within and between networks, we applied a network cohesion analysis as described in Raz *et al.* (2012). The network cohesion index (NCI) is computed continuously in sliding time-windows, based on fMRI data. After a certain set of brain regions is defined, NCI is calculated in a way, which is sensitive to the strength of the average correlation within this set and to the variation about this average. The temporal patterns of intra (within a specific network) and inter (between a pair of networks) NCIs are then compared with behavioral and autonomic indices of empathy.

Continuous retrospective rating of sadness intensity served as a behavioral index for affective empathic engagement. Such operationalization relies on the notion that sadness experienced in response to the observed other's sadness reflects emotional sharing and empathy. Since both films introduce characters in stereotypically sad situations of loss, continuous rating of sadness intensity induced by these unfolding events was regarded as a state-empathy index. Furthermore, in light of evidence associating parasympathetic activity with certain modes of sadness (Kreibig, 2010) and empathic concern (Hastings *et al.*, 2006), a continuous parasympathetic index based on heart rate variability (see 'Materials and Methods' section) was adopted as an autonomic index of empathic reaction. Along with this inquiry of empathy as a transient *state*, the network cohesion indices were also compared with indices of *trait* empathy, as assessed by the commonly used Interpersonal Reactivity Index (IRI; Davis, 1983).

Our hypotheses stand on the assumption that when a certain network increasingly and cohesively engages in the processing of relevant information, the functional connectivity between its nodes will accordingly rise. Relying on the literature on the neural bases of ES and ToM processes, we specifically speculated that the connectivity within these networks, as well as their links to limbic structures, will covary with the intensity of empathy experienced during film viewing. As both of the films induce empathic reactions (i.e. sad responses to other's sadness), we hypothesized positive correlations between the behavioral and physiological indices of empathy and the NCIs computed for the networks of interest specified above.

MATERIALS AND METHODS

Participants

Sixty-four healthy Israelis with at least 12 years of education and no history of neurological or psychiatric disorders were recruited; 21 and 20 data sets for *Stepmom* and *Sophie*, respectively, were discarded due to various technical failures and exaggerated head motions (see below). Thus, 43 (26.93 ± 4.86 years, 22 females) and 44 (26.73 ± 4.69 years, 25 females) valid data sets were available for *Stepmom* and *Sophie*, respectively.

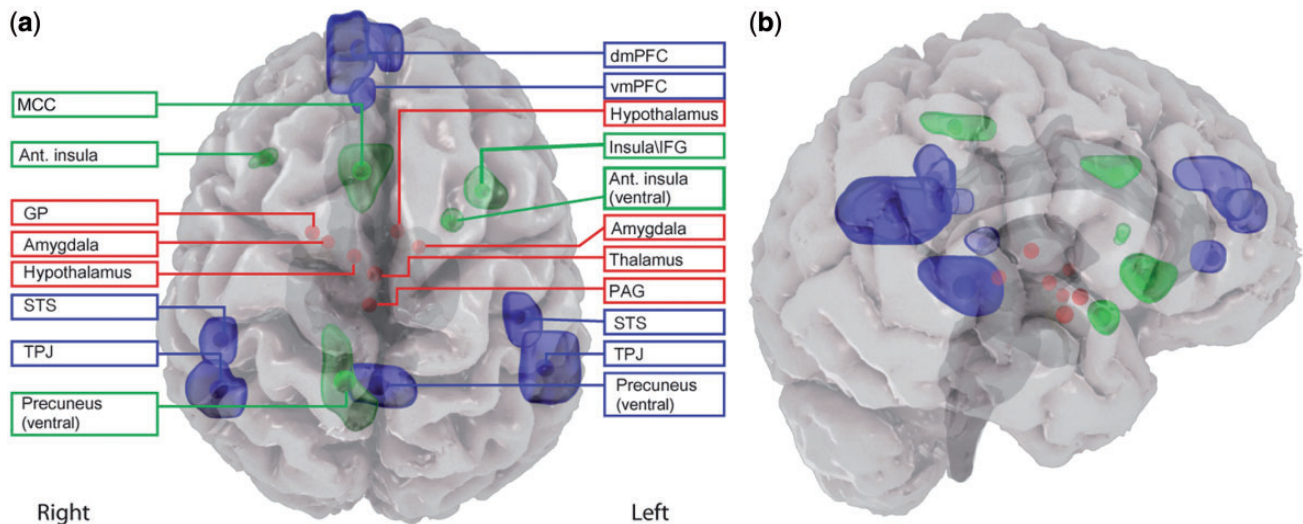


Fig. 1 Locations of ROIs comprising the limbic (red), ES (green) and ToM (blue) networks. Coronal (a) and sagittal (b) views are presented. The ROIs are projected onto a 3D anatomical image adopted from SPL-PNL Brain Atlas (Talos et al., 2008) registered to Talairach space and visualized using a 3D slicer (Pieper et al., 2004) and Autodesk[™] MAYA.

Induction of emotional experience

Two video excerpts from the films *Sophie's Choice* (10 min) and *Stepmom* (8:27 min) were displayed in a counterbalanced order. One-minute long epochs of blank screen were introduced before and after the film presentation. In addition, a 10 min sequence of anatomical scans separated between the clips to reduce emotional carryover.

The excerpt from *Sophie* presents a mother forced by a Nazi officer to choose which of her two children will be sent to death. In the clip from *Stepmom*, a mother talks with both of her children separately about her future death due to a terminal disease.

State-empathy indices

Retrospective emotion rating. In a session conducted immediately following scanning, the participants re-watched the clips while continuously reporting on shifts in sadness intensity during the first viewing in the scanner. The rating was performed retrospectively to avoid distractions during the fMRI recording. This method previously showed good reliability and validity (Raz et al., 2012). The rating was sampled at 10 Hz using homemade software. The participants referred to a vertical scale constantly, indicating seven levels of sadness—from neutral to very deep (each containing 3° of shift; 21° in total).

The parasympathetic index. Electrocardiography (ECG) was obtained during scanning by a BrainAmp ExG MR psychophysiological monitoring system (BrainProducts, Munich). The signal was recorded by Ag/AgCl electrodes at a sampling rate of 500 Hz, yielding 77 data series. A parasympathetic index was computed on the basis of the ECG signal and compared with the behavioral index as described in Raz et al. (2012) and Supplementary S1.

Trait-empathy indices

The IRI is a widely used 28 item self-report multi-dimensional measure of trait-empathy (Davis, 1983) with confirmed validity and reliability (Davis, 1994). It consists of four subscales; indexing personal tendencies to adopt another's point of view on given situations, transposing oneself into the state of fictitious characters, generally experiencing 'other-oriented' feelings and feelings of distress in reaction to the aversive emotions of others. Questionnaires were obtained from 34

(27.04 ± 5.03 years, 16 females) and 34 (26.54 ± 4.49 years, 18 females) participants with valid fMRI data for *Stepmom* and *Sophie*, respectively.

fMRI acquisition

All scans were performed using a GE 3T Signa Excite echo speed scanner with an 8-channel head coil located at the Wohl Institute for Advanced Imaging at the Tel-Aviv Sourasky Medical Center. Structural scans included a T1-weighted 3D axial spoiled gradient echo (SPGR) pulse sequence [time repetition (TR)/TE = 7.92/2.98 ms, flip angle = 15°, pixel size = 1 mm, FOV = 256 × 256 mm, slice thickness = 1 mm]. Functional whole-brain scans were performed in interleaved order with a T2*-weighted gradient echo planar imaging pulse sequence (TR/TE = 3000/35 ms, flip angle = 90°, pixel size = 1.56 mm, FOV = 200 × 200 mm, slice thickness = 3 mm, 39 slices per volume). Active noise cancelling headphones (Optoacoustics) were used.

fMRI data preprocessing

Preprocessing was performed using Brain Voyager QX version 2.1.4. Head motions were detected and corrected using trilinear and sinc interpolations, respectively, applying rigid body transformations with three translation and three rotation parameters. Four data sets were discarded for each of the films due to exaggerated head motions (deviations >1.5 mm and 1.5° from the reference point). The data were smoothed temporally using a linear trend removal with a high pass filter of 0.008 Hz. Spatial smoothing with a 6 mm FWHM kernel was applied. To avoid the confounding effect of fluctuations in the whole-brain blood oxygenation level-dependent (BOLD) signal, for each TR, each voxel was scaled by the global mean at that time point. Anatomical SPGR data were standardized to 1 × 1 × 1 mm and transformed into Talairach space. SPGR images were then manually co-registered with the corresponding functional maps.

Definition of networks of interest

Relevant comprehensive and updated meta-analyses of neuroimaging studies were used to define ES, ToM and limbic networks (Figure 1 and Supplementary Table S1). The ES network was defined on the basis of a recent meta-analysis on empathy for pain (Lamm et al., 2011). Five regions were defined as the nodes of the ES network, derived from

table 2 of Lamm *et al.* (2011). To note, brain structures included in this network are implicated in other forms of emotional contagion (Bruneau *et al.*, 2012). The definition of the ToM network relied on studies in which participants were instructed to infer other's intentions, thoughts and future actions (Bzdok *et al.*, 2012). Seven regions were selected based on further evidence from another study (Dodell-Feder *et al.*, 2011) that they are the only regions to be reliability activated across different ToM tasks. Finally, we tested the interaction of empathy-related networks with subcortical structures implicated in evaluation of the affective value and in the regulation of autonomic reactions. A core limbic set of regions was accordingly defined on the basis of a meta-analysis, which clustered emotion-related brain structured according to their co-activity across studies (Kober *et al.*, 2008). Regions were selected as described in Raz *et al.* (2012), with the exclusion of a left ventral striatal node due to its vicinity to the borders of a ventral insular structure included in the ES network.

In the cases of ES and ToM, the statistical maps were kindly provided by the authors of Lamm *et al.* (2011) and Bzdok *et al.* (2012), respectively. In the case of the limbic network, seed coordinates were obtained from table 4 of Kober *et al.* (2008). MNI to Talairach transformations were performed in the cases of the ES and limbic networks using a Lancaster transformation (Lancaster *et al.*, 2007).

Computation of network cohesion indices

NCI allows a dynamic probing of functional connectivity within and between networks (for details, see Raz *et al.*, 2012). First, the average signal of each ROI was extracted using a Gaussian mask with 3 mm radius around the seed coordinates. Next, for each network, time-window and participant, the set of all pairwise Pearson correlations was computed. The time-windows spanned over 30 s (10 TRs) in line with evidence on dynamic time-dependent patterns of functional connectivity at this time scale (Honey *et al.*, 2007). A right tailed Student's *t*-test with a null hypothesis of $\mu_r = 0$ was performed on the population of the Fisher *Z*-transformed coefficients for each participant, network and time-window. In this test, the *t*-statistic serves as a probe for the connectivity within the network with high values when the mean correlation is high and variance is low.

The intra-NCI for a specific participant, network and time-window was defined as the corresponding *t*-statistic computed as described above. Inter-NCI was calculated in the same manner, except that the population of the *t*-tested pairwise correlations includes pairs of ROIs in different networks.

Comparing NCI with state-empathy indices

The temporal patterns of intra- and inter-NCI were compared with continuous behavioral and physiological indices at the individual level. The first 7 TRs of the cinematic conditions (approximately the span of hemodynamic response) were excluded since they assumingly bear an effect of novelty appreciation related to the onset of the video display. Next, state-empathy indices were down-sampled to fit the temporal framework of the NCIs. Median values of the rated sadness intensity and the parasympathetic index were computed in time-windows of 10 TRs with an overlap of 9 TRs. For each of the participants, Spearman's rank correlation coefficient was computed between these down-sampled indices and the NCIs. These correlations were computed for series including only non-overlapping time-windows of the indices to avoid violation of the assumption of independence between samples in Spearman's test.

For each network or pair of networks of interest, a two-tailed *Z*-test was performed to test whether the average R_s for the comparisons between the NCI and the state-empathy indices across participants

differ from 0. Twenty comparisons were performed at this stage and a false discovery rate (FDR; Benjamini and Hochberg, 1995) correction was applied to control for multiple comparisons. To test the difference between the cinematic conditions in terms of NCI-state-empathy association, *R* populations were compared using a two-tailed paired *t*-test on participants for whom valid fMRI data were available for both films ($N = 34$, 19 females, 26.81 ± 4.81 years).

In addition, to control for a 'whole-brain effect', a spatial bootstrapping approach was employed (see Supplementary S2).

Comparing NCI with trait-empathy indices

This analysis tested the hypothesis that the connectivity of the networks of interest during moments of highly empathic stimulation is modulated by a priori tendency of the participant to be empathically engaged in such states. The time-window that was used to test this hypothesis was selected on the basis of the sadness rating at the group level. We were specifically interested in the dramatic build-up since we expected the relevant networks to be involved in the processing of the highly salient incoming empathy-related information. Thus, the global maxima of the average rating were defined for each of the cinematic conditions. A time-window of 10 TRs coinciding with the TR of maximal rating was selected for each film. For each network or pair of networks of interest, the NCI values in this time-window (after *Z*-scoring the entire NCI signal to standardize individual data) were compared with the IRI scores in each of the four subscales across participants. Spearman's rank test was performed, yielding an overall of 40 comparisons considered for FDR correction.

RESULTS

Relations between behavioral and parasympathetic indices

The comparison of the continuous behavioral and physiological state-empathy indices revealed an opposing pattern between the films. While in *Stepmom* the retrospective sadness rating was positively correlated with the online parasympathetic index ($Z = 3.87$, $P < 0.0002$), in *Sophie* these indices were negatively correlated ($Z = -2.27$, $P < 0.025$; Figure 2). This replicates our previous findings obtained outside of the scanner in a different group of participants (Raz *et al.*, 2012).

Relations between network cohesion and state-empathy indices

Similarly, an opposite pattern of relations between the NCIs and state-empathy indices was evident. The ToM-NCI was *positively* linked with the behavioral index in *Stepmom* [$Z = 2.8$, $P < 0.006$, FDR-adjusted *P*-value (pFDR) < 0.03], but strongly *negatively* correlated with this index in *Sophie* ($Z = -4.34$, $P < 2 \times 10^{-5}$, pFDR < 0.0002 ; Figure 3a). These results showed high specificity in the brain space, as higher and lower behavioral-NCI associations were found for 97.7% and 98.3% of 1000 randomized sets of regions in *Sophie* and *Stepmom*, respectively (Supplementary Figure S2). A direct comparison between the two cinematic conditions confirmed that ToM-NCI was more positively correlated with the sadness rating in *Stepmom* than in *Sophie* ($P < 0.0002$). An association between the ToM-NCI and the parasympathetic index was not found in either case.

At odds with our expectations, the ES-NCI alone did not significantly correlate with indices of state empathy in *Sophie*. Although in this case the peak average ES-NCI and the median sadness rating coincided, with a marginally significant correlation between these variables ($P < 0.08$, uncorrected; *Stepmom*: $P < 0.13$, Supplementary Figure S1), the evidence on their coupling is weak. However, when the relations between ES-limbic-NCI and the rated sadness intensity was tested, a strong positive correlation ($Z = 4.9$, $P < 2 \times 10^{-6}$, pFDR $< 3 \times 10^{-5}$) with high specificity in brain space (99.8%) was found in

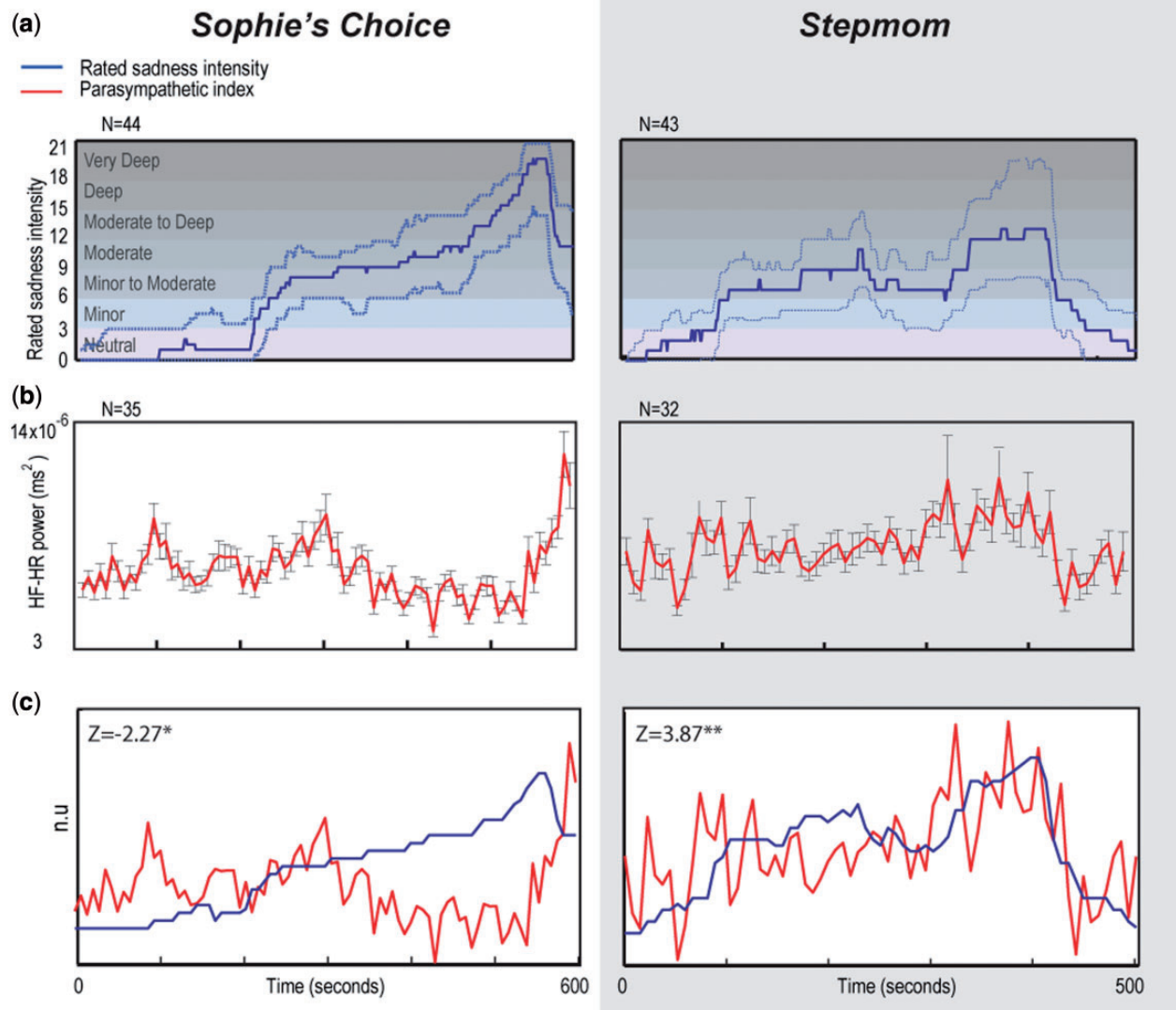


Fig. 2 Behavioral and physiological indices of reactions to film excerpts: (a) rated sadness intensity — median values (solid line) and inter-quartile range (delineated by dashed lines). $N = 44$ and $N = 43$ in *Sophie* and *Stepmom*, respectively. (b) Parasympathetic index (mean \pm standard error) and median rating of sadness intensity resampled as described in the 'Materials and Methods' section. (c) The behavioral and the parasympathetic indices. $N = 35$ and $N = 32$ in *Sophie* and *Stepmom*, respectively. * $P < 0.03$, ** $P < 0.0002$.

Sophie (Figure 3b). This association did not survive the correction for multiple comparisons in *Stepmom*.

ES and ES-limbic-NCIs were not associated with the parasympathetic index. However, the ToM-limbic inter-NCI—although did not correlate with the behavioral index for both films—was nonetheless positively linked with the parasympathetic index ($Z = 2.58$, $P < 0.01$, pFDR < 0.04 , specificity: 92.17%) in *Stepmom* (Figure 3d).

Markedly, the limbic-NCI was *positively* correlated with the rated sadness intensity ($Z = 4.33$, $P < 2 \times 10^{-5}$, pFDR < 0.001 , specificity: 99.5%; Supplementary Figure S1), but *negatively* correlated with the parasympathetic index ($Z = -2.95$, $P < 0.004$, pFDR < 0.02 , specificity: 99.6%; Figure 3d) in *Sophie*. No such associations were found in *Stepmom*.

Relations between network cohesion and trait-empathy indices

We next tested the hypothesis that personal empathic tendencies influence the connectivity of empathy-related networks during moments of high emotional engagement. Peak emotional intervals were defined on the basis of the average sadness rating. NCI for these intervals were

then compared with the scores of the four IRI subscales for each of the films.

During these intervals of peak empathic engagement, ToM intra-NCI positively correlated with the personal distress scores ($R_s = 0.53$, $P < 0.002$, pFDR < 0.03) in *Stepmom*, whereas in *Sophie* the same subscale scores were associated with ES-limbic inter-NCI ($R_s = 0.53$, $P < 0.002$, pFDR < 0.03). No other correlation between the NCIs and the trait-empathy subscale scores survived FDR correction.

DISCUSSION

This study provides unique evidence regarding the neural bases of empathic experience, both corroborating and potentially refining the accumulated knowledge on this topic. Our findings augment converging data on distinct empathy-related processes, relying on either insular-cingulate or MPF-temporo-parietal circuits (associated with ES and ToM, respectively). Particularly—and to our knowledge, unprecedentedly—we found the *dynamic* patterns of *connectivity* of these circuits to be associated with empathy experienced under realistic

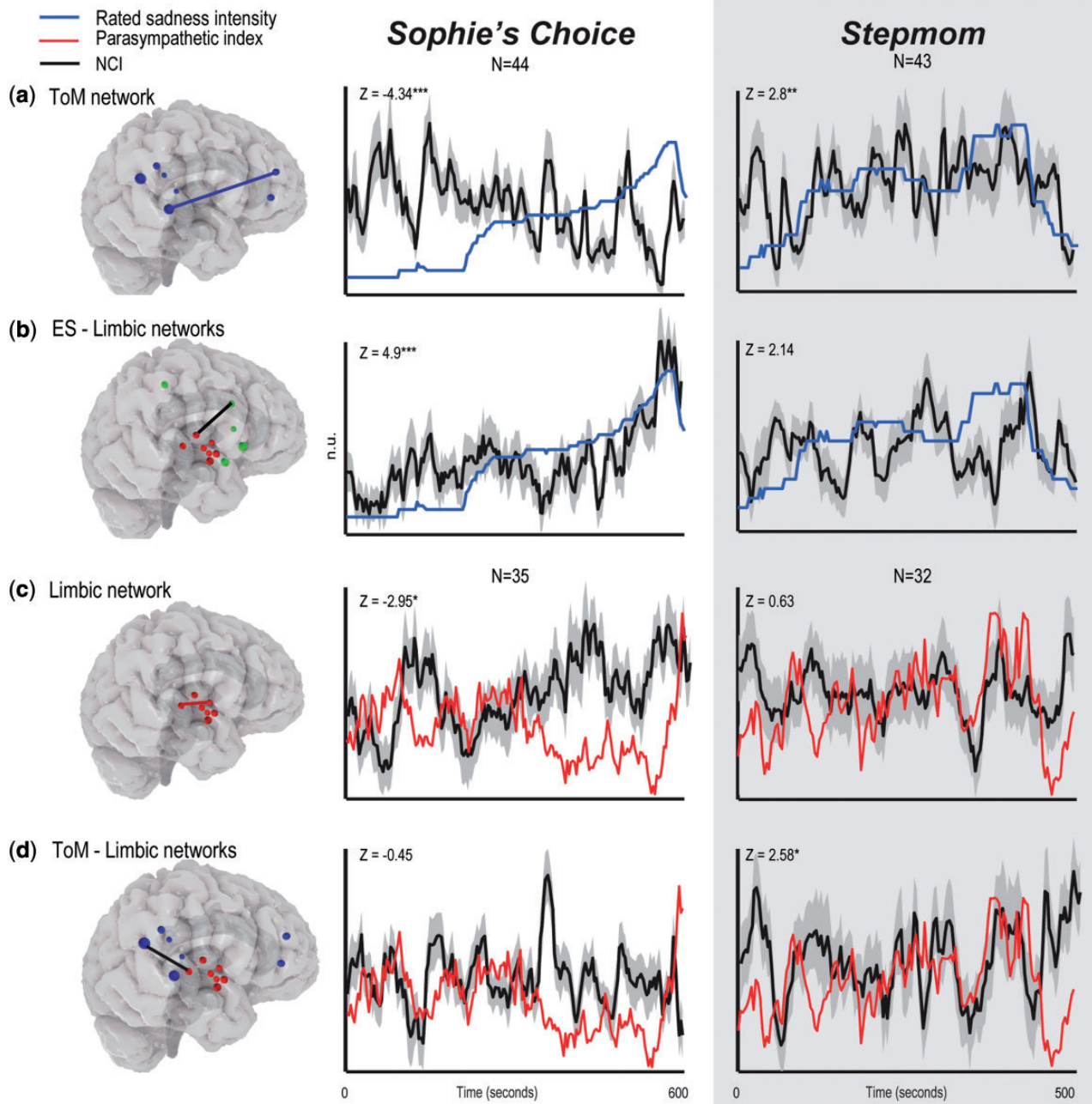


Fig. 3 Relations between behavioral, physiological and network cohesion indices. Median values of the behavioral index (a,b) and mean values of the physiological (c,d) and network cohesion (a–d) indices are indicated. The gray surface indicates ES-NCI standard error. $N = 44$ and $N = 43$ in *Sophie* and *Stepmom*, respectively. * $P < 0.01$, $pFDR < 0.04$, ** $P < 0.006$, $pFDR < 0.03$, *** $P < 2 \times 10^{-5}$, $pFDR < 0.001$.

situations. Furthermore, our data indicate a growing interaction of these circuits with a set of subcortical limbic structures during the intensification of empathic engagement. However, these findings also evince a context-dependent dissociation between empathy-related brain processes, suggesting that emotional sharing is based on the interplay between ES- or ToM-related processes, which may alternatively dominate empathic engagement.

Different empathy-related systems were found to be associated with empathic engagement for each of the films. In *Stepmom*, it was the ToM-NCI that was correlated both with the state-empathy index (sadness rating) over time, and the trait-empathy index (personal distress scores) in a selected time-window of peak rated sadness. Strikingly, these same state and trait indices were strongly associated with ES-

limbic-NCI in *Sophie*. Thus, in each of the cinematic conditions, it was the NCI that most significantly positively correlated with the state-empathy behavioral index that also correlated with the personal distress scores in the tested time-window.

How should we interpret the indications that it is specifically the interaction between these subcortical and cingulo-insular structures that underpins aware empathic experience in *Sophie*? Notably, these findings are congruent with Craig’s model of the neural basis of emotional experience (Craig, 2009). This model posits that the real-time experience of a sense of self emerges as a result of an interaction between brain structures within both the ES and the limbic networks. According to this hierarchical model, interoceptive information about the homeostatic state of the body ascends from the spinal cord to

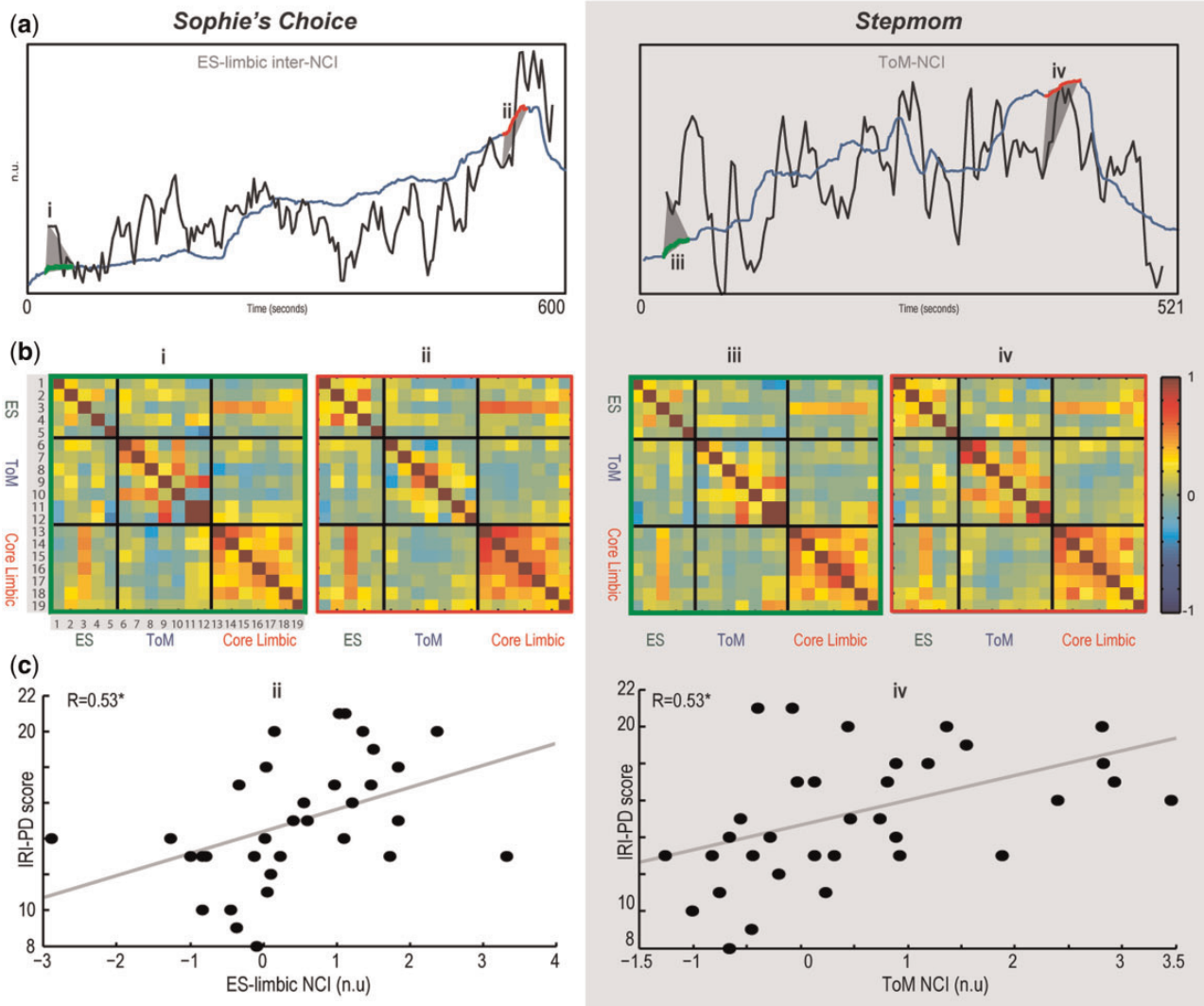


Fig. 4 (a) Cohesion indices of selected networks (gray) and non-resampled average rating. Time-windows of culminating reported empathic engagement are marked in red. (b) Heat maps representing the correlation matrices used for NCI computation. The right map in each pair relates to the selected time-window (see a), whereas the left map refers to the first time-window for comparison. For each pair of ROIs, the average correlation over participants is indicated by a color. For ROIs nomenclature, see Figure 1. (c) NCI values of the selected networks in the time-windows of interest against trait-empathy index of IRI-PD ($N = 34$).

specific thalamic nuclei and is then projected to the insular and anterior cingulate cortices (ACC). The AI then integrates multimodal information from various sources including affective information coming from relatively low-level limbic areas such as the hypothalamus and the amygdala. Subsequently, a motivational value for the integrated representations of the interoceptive state is established and fed back to the body via major descending projection from the ACC to the PAG. Thus, when another's distress is empathically experienced as one's own interoceptive state, the interaction between the components of this circuit (reflected as ES-limbic-NCI) is indeed expected to increase.

The indications for the involvement of this ES-limbic circuit were weaker in the case of *Stepmom*, for which evidence on ToM-related processing was found. Importantly, evidence suggests that the recruitment of ToM- and ES-related neural circuits varies as a function of the temporal relations between the subject and the object of mental processing, as well as of the interpersonal distance between the two. Thus, Buckner and Carroll (2007) emphasize the role of MPF and medial temporal structures in processes that rely on the ability to flexibly shift

one's perspective beyond the immediate present in contrast with processing bound to real-time input, which recruits ES and limbic regions (Craig, 2009). These authors evince that ToM considerably shares its neural correlates with other cognitive functions such as autobiographical memory and thinking about the future that involve similar detachment from the here-and-now perspective.

Moreover, in terms of *agency*, the insula and the ACC have been associated with 'first-person' simulations of other's state involving minimal distance between individuals, whereas the STS, TPJ and mPFC have been implicated with the projection of one's own properties to an objectified, distant or separated agent. When the perceiver of animated or filmed actions allegedly believed that she controls these actions, insular activations were found, but when these actions were perceived as controlled by another agent, temporo-parietal regions were recruited (e.g. Farrer and Frith, 2002; Corradi-Dell'acqua *et al.*, 2008). Other evidence suggests that direct stimulation of TPJ may even induce an extreme 'third-person' perspective in which individuals report that they see their own bodies as separated external observers (e.g. Blanke *et al.*, 2002). A highly relevant recent study of both agency

and empathy reports that witnessing the social exclusion of a *stranger* is associated with the activation of the dorsal mPFC, while dorsal ACC and AI are activated in response to the exclusion of *close friend* proportionally to the strength of their tendency to share emotions with that friend (Meyer *et al.*, 2012).

Does this distinction between the modes of agency correspond to a difference between the contents of *Sophie* and *Stepmom*? A systematic examination of the cinematic factors that induce increased ToM-related processing in *Stepmom* and ES-limbic integration in *Sophie* is yet to be conducted. However, a key thematic distinction between the clips, which is related to agency, may readily be considered relevant: both films introduce a theme of separation of mother from child, but in *Sophie* the loss is presented as a real-time probabilistic event whereas in *Stepmom* the loss is presented as a determined fact, which cannot be changed by intentional action. In *Stepmom*, the mother and children discuss the separation as a given fact, while *Sophie* and her children face an unfolding act of separation. Therefore, it is possible that *Sophie* triggers a ‘first-person engagement’ wherein the film viewer and the cinematic characters share an increasingly integrated activity of the ES-limbic circuit responsible for viscerally based sensations during a real-time action. On the other hand, in *Stepmom*, when the loss is primarily simulated as a distant and objective event, ToM-related processing, facilitating a flexible representation of non-actual states, may mediate empathic engagement.

A support for this interpretation comes from a recent meta-analysis of emotion-related autonomic reactions (Kreibig, 2010). Reviewing multiple instances of experimentally induced sadness, two major modes of autonomic reaction were identified, involving either increased sympathetic activity or enhanced parasympathetic control. Importantly, this meta-analysis reports that an ‘activating’ mode of sadness response appears predominantly in reaction to films that present loss as an event that is expected to occur at any moment, as in *Sophie*, whereas a ‘deactivating’ parasympathetic pattern characterizes responses to cinematic representation of loss as a determined fact (Kreibig, 2010), as in *Stepmom*.

Correspondingly, our findings indicate that *Stepmom* elicits an experience of heightened parasympathetic tone as reflected by an association between the sadness rating and the autonomic index. This parasympathetic reaction is specifically associated with the processing of affective aspects of ToM information as indicated by the correlation between ToM-limbic-NCI and the parasympathetic index. On the other hand, the emotional intensification in *Sophie* is related to decreased parasympathetic activity, evident as negative correlations between the parasympathetic index and sadness rating, as well as the limbic-NCI. Taken together, this line of evidence suggests that the alternative recruitment of ToM- and ES-related circuits in empathic situations is linked with the adoption of either a ‘third-person’ perspective on separated agents (Tan, 1995) or a proactive ‘first-person’ involvement with an online situation.

However, one key finding remains surprising also within the suggested explanatory framework. Assuming that a global drop in the correlations between the nodes of a certain network reflects the attenuation of its function as an integrative module, the apparent drop in ToM-NCI during the most dramatic moments in *Sophie* has yet to be explained. Why this empathy-related mode of processing is released or even suppressed in the presence of other’s extreme distress?

Notably, mentalization and ES are described in the literature as both complementary and alternative modes of empathy (Jankowiak-Siuda *et al.*, 2011). ES may effectively motivate pro-social behavior and thus have a key value for survival. However, if the self/other distinction is substantially undermined, a risk is posed to one’s ability to function autonomously according to non-empathic individual needs. Maintaining a distance from the object of empathy may therefore

‘protect’ the empathizer from automatic and maladaptive contagious reactions (Jankowiak-Siuda *et al.*, 2011). This ability, which develops with age (Hastings *et al.*, 2006), helps to regulate empathic distress as it relies on a separation between one’s own and other’s perspectives of a given situation, implying that suffering is not self-originated. In fact, the ability to generate this differentiation by conceiving of the others as separate agents with distinct concerns, beliefs and intentions is a developmental precondition for gaining the ability to engage in such ToM processing.

It is possible that the dip in ToM intra-NCI, while viewing *Sophie*’s imposed choice to send her daughter to her death, reflects the suspension of this regulated and distanced form of cognitive engagement (allowing ‘feeling for’) in favor of an embodied and direct form of emotional sharing (‘feeling with’). Indeed, in line with the notion of regulatory function of ToM, MPF regions are often implicated in emotion regulation as well (e.g. Goldin *et al.*, 2008). Moreover, the evidence on the role of the parasympathetic system in emotion regulation (Porges, 1991; Rottenberg *et al.*, 2003) together with its relevance to the emotional experience elicited by *Stepmom* indicates the regulatory aspect of the mode of empathic engagement elicited by this film. It should be noted that the part of the ToM-limbic circuit that apparently increased its connectivity hand-in-hand with parasympathetic activation has been implicated in the regulation of efferent outflow both through direct projection and via indirect relay pathways (Dampney, 2011).

In conclusion, relying on a novel approach for analyzing the dynamics of functional network connectivity, this study provides further support for the claims of the involvement of cingulo-insular and MPF-temporo-parietal regions in empathy-related processes of ES and ToM, respectively. The inter-relations between the multi-level indices suggest that *Stepmom* and *Sophie* comprise instances of two forms of empathy for sadness. The first relies on a cognitive representation of another’s distress from a ‘third-person’ perspective facilitating regulated reaction. The second involves a ‘first-person’ proactive experience of automatic emotional resonance, accompanied by the attenuation of a more psychologically distanced judgment of the situation. The network analysis of ES and ToM circuits specifically suggests that (i) the contribution of these circuits to the generation of empathic reactions is context-dependent and may be mediated by their interactions with subcortical limbic structures. These circuits may be sinusously recruited depending on the circumstances under which empathy is experienced. (ii) ToM-related processing in particular may be enhanced but also diminished when confronting another’s distress. This phenomenon may reflect a flexible switching between more or less cognitively regulated modes of engagement.

SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

Conflict of Interest

None declared.

REFERENCES

- Baron-Cohen, S., Leslie, A.M., Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21, 37–46.
- Barrett, L.F., Mesquita, B., Ochsner, K.N., Gross, J.J. (2007). The experience of emotion. *Annual Review of Psychology*, 58, 373.
- Batson, C.D., Duncan, B.D., Ackerman, P., Buckley, T., Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, 40, 290–302.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57, 289–300.

- Blanke, O., Ortigue, S., Landis, T., Seeck, M. (2002). Stimulating illusory own-body perceptions. *Nature*, *419*, 269–70.
- Bruneau, E.G., Pluta, A., Saxe, R. (2012). Distinct roles of the “Shared Pain” and “Theory of Mind” networks in processing others’ emotional suffering. *Neuropsychologia*, *50*, 219–31.
- Buckner, R.L., Carroll, D.C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*, 49–57.
- 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 and Function*, *217*, 783–96.
- Corradi-Dell’acqua, C., Ueno, K., Ogawa, A., Cheng, K., Rumiati, R.I., Iriki, A. (2008). Effects of shifting perspective of the self: an fMRI study. *Neuroimage*, *40*, 1902–11.
- Craig, A.D. (2008). Interoception and emotion: a neuroanatomical perspective. In: Lewis, M., Haviland-Jones, J., editors. *Handbook of Emotions*. New York: The Guilford Press, pp. 272–88.
- Craig, A.D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews. Neuroscience*, *10*, 59–70.
- Dampney, R.A.L. (2011). The hypothalamus and autonomic regulation: an overview. In: Llewellyn-Smith, I.J., Verberne, A.J.M., editors. *Central Regulation of Autonomic Functions*. New York: Oxford University Press, pp. 47–61.
- Danziger, N., Faillelot, I., Peyron, R. (2009). Can we share a pain we never felt? Neural correlates of empathy in patients with congenital insensitivity to pain. *Neuron*, *61*, 203–12.
- Davis, M.H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–26.
- Davis, M.H. (1994). *Empathy: A Social Psychological Approach*. Social Psychology Series. Boulder, CO: Westview Press.
- De Waal, F.B.M. (2008). Putting the Altruism back into Altruism: the evolution of empathy. *Annual Review of Psychology*, *59*, 279–300.
- Decety, J. (2011). The neuroevolution of empathy. *Annals of the New York Academy of Sciences*, *1231*, 35–45.
- Dodell-Feder, D., Koster-Hale, J., Bedny, M., Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, *55*, 705–12.
- Eisenberg, N., Strayer, J. (1987). Critical issues in the study of empathy. In: Eisenberg, N., Strayer, J., editors. *Empathy and its Development*. Cambridge: Cambridge University Press, pp. 3–13.
- Eisenberger, N.I., Lieberman, M.D., Williams, K.D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, *302*, 290–2.
- Farrer, C., Frith, C.D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage*, *15*, 596–603.
- Gallagher, H.L., Frith, C.D. (2003). Functional imaging of “theory of mind”. *Trends in Cognitive Sciences*, *7*, 77–83.
- Gallese, V. (2005). Embodied simulation: from neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, *4*, 23–48.
- Gallese, V. (2007). Before and below “theory of mind”: embodied simulation and the neural correlates of social cognition. *Philosophical transactions of the Royal Society of London. Series B*, *362*, 659–69.
- Goldin, P.R., Hutcherson, C.A., Ochsner, K.N., Glover, G.H., Gabrieli, J.D., Gross, J.J. (2005). The neural bases of amusement and sadness: a comparison of block contrast and subject-specific emotion intensity regression approaches. *Neuroimage*, *27*, 26–36.
- Goldin, P.R., McRae, K., Ramel, W., Gross, J.J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*, 577–86.
- Harrison, N.A., Singer, T., Rotshtein, P., Dolan, R.J., Critchley, H.D. (2006). Pupillary contagion: central mechanisms engaged in sadness processing. *Social Cognitive and Affective Neuroscience*, *1*, 5–17.
- Hastings, P.D., Zahn-Waxler, C., McShane, K. (2006). We are, by nature, moral creatures: biological bases of concern for others. In: Killen, M., Smetana, J.G., editors. *Handbook of Moral Development*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 483–516.
- Honey, C.J., Kötter, R., Breakspear, M., Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 10240–5.
- Jabbi, M., Bastiaansen, J., Keysers, C. (2008). A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways. *PLoS One*, *3*, e2939.
- Jankowiak-Siuda, K., Rymarczyk, K., Grabowska, A. (2011). How we empathize with others: a neurobiological perspective. *Medical Science Monitor*, *17*, RA18–24.
- Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D. (2008). Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage*, *42*, 998–1031.
- Krämer, U.M., Mohammadi, B., Doñamayor, N., Samii, A., Münte, T.F. (2010). Emotional and cognitive aspects of empathy and their relation to social cognition—an fMRI-study. *Brain Research*, *1311*, 110–20.
- Kreibig, S.D. (2010). Autonomic nervous system activity in emotion: a review. *Biological Psychology*, *84*, 394–421.
- Lamm, C., Decety, J., Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*, 2492–502.
- Lancaster, J.L., Tordesillas-Gutiérrez, D., Martínez, M., et al. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping*, *28*, 1194–205.
- Lazarus, R.S. (1994). *Emotion and Adaptation*. New York: Oxford University Press.
- Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F. (2012). The brain basis of emotion: a meta-analytic review. *The Behavioral and Brain Sciences*, *35*, 121–43.
- Lombardo, M.V., Chakrabarti, B., Bullmore, E.T., et al. (2009). Shared neural circuits for mentalizing about the self and others. *Journal of Cognitive Neuroscience*, *22*, 1623–35.
- Meyer, M.L., Masten, C.L., Ma, Y., et al. (2012). Empathy for the social suffering of friends and strangers recruits distinct patterns of brain activation. *Social Cognitive and Affective Neuroscience*, *8*, 446–54.
- Oatley, K.G. (1996). Emotions, rationality and informal reasoning. Mental models in cognitive science: Essays in honour of Phil Johnson-Laird, 175–196.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews. Neuroscience*, *9*, 148–58.
- Pieper, S., Halle, M., Kikinis, R. (2004). 3D Slicer. *Proceedings of the IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2004, Arlington, VA, pp. 632–5.
- Porges, S.W. (1991). Vagal tone: an autonomic mediator of affect. In: Garber, J., Dodge, K.A., editors. *The development of emotion regulation and dysregulation. Cambridge studies in social and emotional development*. New York: Cambridge University Press, pp. 111–28.
- Raz, G., Winetraub, Y., Jacob, Y., et al. (2012). Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. *NeuroImage*, *60*, 1448–61.
- Rottenberg, J., Wilhelm, F.H., Gross, J.J., Gotlib, I.H. (2003). Vagal rebound during resolution of tearful crying among depressed and nondepressed individuals. *Psychophysiology*, *40*, 1–6.
- Russell, J.A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion*, *23*, 1259–83.
- Scherer, K.R. (2009). The dynamic architecture of emotion: evidence for the component process model. *Cognition & Emotion*, *23*, 1307–51.
- Shamay-Tsoory, S.G. (2011). The neural bases for empathy. *Neuroscientist*, *17*, 18–24.
- Smith, M. (1995). *Engaging characters: fiction, emotion, and the cinema*. Oxford, England: Clarendon.
- Talos, I.-F., Jakab, M., Kikinis, R., Shenton, M. (2008). SPL-PNL Brain Atlas (WWW Document) <http://www.slicer.org/publications/item/view/1265>.
- Tan, E.S. (1995). *Emotion and the Structure of Narrative Film: Film as an Emotion Machine*. Mahwah, NJ: Erlbaum.
- Van Overwalle, F., Baetens, K. (2009). Understanding others’ actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, *48*, 564–84.
- Wicker, B., Keysers, C., Plailly, J., Royet, J.-P., Gallese, V., Rizzolatti, G. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–64.
- Zaki, J., Ochsner, K. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature Neuroscience*, *15*, 675–80.