



## UvA-DARE (Digital Academic Repository)

### Changes in brain activity following the voluntary control of empathy

Borja Jimenez, K.C.; Abdelgabar, A.R.; De Angelis, L.; McKay, L.S.; Keysers, C.; Gazzola, V.

**DOI**

[10.1016/j.neuroimage.2020.116529](https://doi.org/10.1016/j.neuroimage.2020.116529)

**Publication date**

2020

**Document Version**

Final published version

**Published in**

NeuroImage

**License**

CC BY-NC-ND

[Link to publication](#)

**Citation for published version (APA):**

Borja Jimenez, K. C., Abdelgabar, A. R., De Angelis, L., McKay, L. S., Keysers, C., & Gazzola, V. (2020). Changes in brain activity following the voluntary control of empathy. *NeuroImage*, 216, [116529]. <https://doi.org/10.1016/j.neuroimage.2020.116529>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



## Changes in brain activity following the voluntary control of empathy

K.C. Borja Jimenez<sup>a,1</sup>, A.R. Abdelgabar<sup>a,1</sup>, L. De Angelis<sup>a,1</sup>, L.S. McKay<sup>a,c,1</sup>, C. Keysers<sup>a,b,2</sup>, V. Gazzola<sup>a,b,\*,1</sup>

<sup>a</sup> Social Brain Lab, Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105 BA, Amsterdam, the Netherlands

<sup>b</sup> Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 166, 1018, WV, Amsterdam, the Netherlands

<sup>c</sup> Current Address: Division of Psychology, School of Education & Social Sciences, University of the West of Scotland, High Street, Paisley, PA1 2BE, UK

### ARTICLE INFO

#### Keywords:

Cognitive control  
Emotion regulation  
Intersubject correlation  
Functional connectivity  
Reappraisal  
Theory of mind

### ABSTRACT

In neuroscience, empathy is often conceived as relatively automatic. The voluntary control that people can exert on brain mechanisms that map the emotions of others onto our own emotions has received comparatively less attention. Here, we therefore measured brain activity while participants watched emotional Hollywood movies under two different instructions: to rate the main characters' emotions by empathizing with them, or to do so while keeping a detached perspective. We found that participants yielded highly consistent and similar ratings of emotions under both conditions. Using intersubject correlation-based analyses we found that, when encouraged to empathize, participants' brain activity in limbic (including cingulate and putamen) and somatomotor regions (including premotor, SI and SII) synchronized more during the movie than when encouraged to detach. Using intersubject functional connectivity we found that comparing the empathic and detached perspectives revealed widespread increases in functional connectivity between large scale networks. Our findings contribute to the increasing awareness that we have voluntary control over the neural mechanisms through which we process the emotions of others.

### 1. Introduction

It is well established that observing another person experiencing emotions triggers representations in brain regions associated with our own actions, sensations, and emotions (Engen and Singer, 2013; Keysers et al., 2010; Keysers and Gazzola, 2009; Lamm et al., 2011). The recruitment of neural representations tied to our own body (in somatosensory and motor regions) and to our own affect (in limbic regions such as the cingulate, insula, striatum, and amygdala) is conceived as an embodied process. This process is associated with the concept of empathy, i.e. experiencing what other people feel while being aware that this vicarious state is produced by someone else (Keysers and Gazzola, 2009; Lamm et al., 2011). This phenomenon can be studied using block designs in which participants either view or experience certain emotions, such as pleasure, disgust, or pain (Jabbi et al., 2007; Keysers et al., 2004; Singer et al., 2004; Wicker et al., 2003). An interesting alternative is to show longer movies that include a variety of emotions, and to identify

those voxels in which activity is synchronized across viewers of the same stimuli (Hasson et al., 2004, 2010; Nummenmaa et al., 2012; Nastase et al., 2019). This approach leverages the fact that for a voxel to synchronize across viewers of the same movie, its activity fluctuations must carry information about the movie, allowing us to map how much information each voxel in the brain has about the content of the movie while using complex stimuli that situate human interactions in context (Nastase et al., 2019).

In addition to embodied processes, ample evidence shows that participants can infer the cognitive mental states of others in more abstract ways using brain regions that include the temporoparietal junction (TPJ), the precuneus, and the medial prefrontal cortex (mPFC; Mar, 2011; Frith and Frith, 2008; Schurz et al., 2014). This process, often called mentalizing (Frith and Frith, 2008) or cognitive empathy (Preston and De Waal, 2002), is typically studied in the context of attributing false beliefs to other individuals (Saxe and Kanwisher, 2003), but it also seems to be involved while we process the emotions of others (Schnell et al., 2011).

\* Corresponding author. Social Brain Laboratory, Netherlands Institute for Neuroscience Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105, BA Amsterdam, the Netherlands.

E-mail address: [v.gazzola@nin.knaw.nl](mailto:v.gazzola@nin.knaw.nl) (V. Gazzola).

<sup>1</sup> Shared first authorship.

<sup>2</sup> Shared last authorship.

<https://doi.org/10.1016/j.neuroimage.2020.116529>

Received 19 June 2019; Received in revised form 19 December 2019; Accepted 7 January 2020

Available online 10 January 2020

1053-8119/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Hence, there seem to be two complementary pathways to perceiving what goes on in others: a cognitive pathway involved in mentalizing, and an embodied pathway often associated with empathy (Keysers and Gazzola, 2007; Schnell et al., 2011).

While mentalizing is typically considered to be relatively voluntary, embodied processes through which witnessing the emotions of others triggers our own emotions are typically regarded as automatic, albeit occurring to varying extent depending on individual differences in trait empathy (e.g. Jabbi et al., 2007; Singer et al., 2004), psychopathology (e.g. Bird et al., 2010; Meffert et al., 2013), or context (e.g. Azevedo et al., 2013; Hein et al., 2010; Singer et al., 2006). Thus, the voluntary control of the embodied processes have been overlooked, with the exception of three notable studies. Lahnakoski et al. (2014) showed that while watching a soap opera people could voluntarily change their perspective from that of a detective, (i.e. paying attention to people), to that of an interior decorator (i.e. paying attention to objects), and that this perspective switch led to measurable changes in how parietal and visual cortices represented the clip. Bruneau et al. (2013) showed that actively trying to empathize with another's emotional pain led to increases in amygdala activity. Meffert et al. (2013) showed that instructions to empathize boosted activity in the insula and the anterior cingulate cortex while witnessing other people in pain.

In contrast to this limited number of neuroscience studies, the psychology literature has provided ample evidence that, besides self-regulation of emotions (Gu and Han, 2007; Lamm et al., 2010; Hallam et al., 2014), people are highly selective in the voluntary control of affective empathy, boosting it when they expect empathy to benefit them, and reducing it when they expect it to be detrimental (see Weisz and Zaki, 2018; Zaki, 2014; Schumann et al., 2014 for reviews, and Cameron et al., 2016; Cikara et al., 2014; Morelli et al., 2015; Pickett et al., 2004; Shaw et al., 1994 as examples of studies using this approach). We would thus expect to find a strong voluntary modulation of brain activity in networks mapping the actions and emotions of others onto our own (Keysers and Gazzola, 2014), akin to the well documented deliberate neural modulations of our own emotions (for a review see Kohn et al., 2014).

The aim of the current experiment was thus to investigate whether and where in the brain we can voluntarily modulate the route through which we attribute emotions to others. For this aim, we asked participants, in an MRI scanner, to watch extracts from Hollywood movies in which the main character undergoes strong emotional fluctuations, and to simultaneously rate the emotional state of the main protagonist from moment to moment. To localize the neural correlates of this voluntary control, people performed the emotion rating task under two different instructions. During the *Empathic* session, participants were required to watch the videos with specific instructions to actively empathize with and share the feelings of the main character while rating them. During the *Detached* session, participants watched the videos with specific instructions to be as detached as possible from the main character in the film, and to try not to share his feelings. In both conditions participants therefore had to continuously rate and perceive the affective state of the main character but while in one case they could let these emotions permeate their own state, in the other, they could have used more cognitive routes without being affected themselves. Based on the differentiation of cognitive and affective empathy we would expect the Empathic manipulation to increase the degree to which brain regions associated with embodied processing synchronize across participants, and hence represent the movie (Nastase et al., 2019) and the Detached instruction to increase synchrony in regions associated with mentalizing.

Having to provide an explicit moment to moment report of other people's emotions undoubtedly influences brain activity relative to a more natural situation in which we react to the emotions of others more implicitly. However, we chose this task to ensure that participants process the emotions of the main character under both conditions, and that we can quantify the accuracy of this processing across conditions using these ratings. Otherwise, differences between the empathic and detached

condition could simply reflect differences between attending and not attending to the emotions of the protagonist (as in Lahnakoski et al., 2014). Given the important distinction between taking a first versus third person perspective (Decety and Meyer, 2008; Reniers et al., 2014), we also explicitly instruct participants in both conditions to rate how the protagonist is feeling (Schnell et al., 2011), not how the participants themselves feel.

After exploring whether instructions alter how participants rate the emotions of the protagonists, we first use intersubject correlation (ISC; Nastase et al., 2019) to examine which voxels are influenced by the change in instructions. Second, we explore intersubject functional connectivity (ISFC; Nastase et al., 2019) to examine whether instructions reshaped the functional connectivity across the brain networks that process the movie.

## 2. Methods

### 2.1. Participants

Twenty-three healthy, right-handed, native English speaking males were recruited from the city of Amsterdam, the Netherlands. Individuals with a history of neurological/psychiatric symptoms or medication usage were excluded from participation. One participant was excluded due to potentially abnormal brain anatomy and their data referred to a neurologist for further examination, and three more participants were excluded due to incomplete data acquisition. An additional participation criterion was being naive to the stimulus material (see section 2.3). Nineteen participants were included in the final sample (mean age = 33.2 years  $\pm$  10.9 SD). All subjects were compensated for their time and travel costs, and gave their informed consent. The study was approved by the Ethics Review Board Committee of the University of Amsterdam under the protocol NL45843.018.

### 2.2. Experimental task

Participants underwent three fMRI sessions on different days in which participants watched the same two movie clips under three task instructions. Only one instruction was given within one session, and the order of instruction was randomized between participants. Participants were informed about the three viewing instructions at the beginning of the experiment (see Online Supplementary Material 1), and reminded at the beginning of each scanning session before the movie presentation.

During the *Empathic* session, participants were required to watch the videos under the following instructions: "we ask that you try to empathize with the target character while making the ratings. By this we mean for you to try to be compassionate, caring and warm towards the character, perhaps even sharing what they are feeling." This instruction was repeated on the screen just before the movie presentation in the following way: "Watch the following video and rate the emotions of the character whilst being empathic, try to be compassionate, caring, warm and compassionate".

During the *Detached* session, participants watched the videos under the following reminder instructions: "During one viewing we ask that you try to be as detached from the target character as possible. During this viewing try not to feel with the character, but rate their emotions from a detached, dispassionate and objective point of view." Just before the beginning of the movie, participants were reminded of the instruction with the text: "Please watch the following video and rate the emotions of the character whilst being detached, objective and dispassionate".

During the *Own* session, participants had to rate their own emotional response to the movie: "During this viewing, we ask that you rate how you are feeling at each moment in time without paying specific attention to a target in the film. Indeed, pay attention to your own emotions during this viewing and how the movie makes you feel. Try to be as honest as you can and remember that the data is anonymised, so we will not be able to share information about you with anyone else." Before the beginning

of the movie their reminder read: “Please watch the following video and rate your own emotions as positive or negative throughout the clip”.

During the movie presentation, participants had to rate the main characters’ emotions (in the Detached and Empathic conditions) or their own (in the Own condition) moment by moment, using a continuous visual analogue scale displayed below the movie with a cursor moved by an fMRI compatible trackball (HHSC-TRK-2, Philadelphia, PA, Current Designs, [www.curdes.com](http://www.curdes.com)). Visually, the scale ranged from ‘very negative’ to ‘very positive’, corresponding to a continuous trackball x-coordinate scale coded from  $-500$  to  $500$ , respectively. To minimize the influence that the clip content could have on the modulation process at the brain and behavioural level, the order of the sessions was randomized between participants, and the clip presentation order counterbalanced across sessions.

As we unfortunately did not include the same instruction manipulation in the self-rating condition (i.e. to rate how the participant feels while being empathic with the protagonist vs. while being detached), and because we did not direct participants’ attention to the protagonist in the movie, it is impossible to know how participants positioned themselves relative to the protagonist while rating their own emotions. The direct comparison of our two main conditions reporting how the protagonists feel vs. the self-rating condition cannot therefore speak to the Empathic/Detached distinction at the core of this manuscript. Additionally, as the emotions felt by participants are likely induced by the movie, this condition cannot be considered as a clean map of the participant’s self-experience of emotions. This condition is therefore not discussed further in this paper. The presence of the Own condition though helped participants to understand that, by contrast, during the Empathic and Detached instruction, they had to focus on the emotions of the main character, rather than their own.

Finally, at the end of the experiment, empathy trait measures (Interpersonal Reactivity Index scores, IRI; Davis, 1980) were also collected for 17 out of the 19 participants. The power to detect significant IRI-ISC correlations is unfortunately very low and running such correlations would lead to inflated estimates and difficult to replicate results (Cremers et al., 2017). Accordingly, we used the IRI only for some exploratory analyses that are reported as Figs. S1–S3 and Table S1.

### 2.3. Stimuli

Naturalistic stimuli are known to be more robust in eliciting emotional responses (Westermann et al., 1996). We have therefore used Hollywood movies to present a naturalistic scenario that portrays the complexity of affective human interactions (see Raz et al., 2012; Raz and Hendlar, 2014; Raz et al., 2016; Vanderwal et al., 2015; Dayan et al., 2018; Nanni et al., 2018 as examples of studies), and thus promote empathic feelings and measurable brain responses (i.e. Hasson et al., 2010; Hasson et al., 2012). As we were interested in modulatory processes that are independent from a specific context or emotion, we decided to present two, rather than one, movie clips extracted from popular dramatic movies that are known to depict and elicit positive and negative emotions. Stimuli clips were extracted from ‘The Champ’ (Lovell, 1979), a drama film reported to induce, amongst others, deep sadness in viewers (Gross and Levenson, 1995; duration 11 min. 46 sec.), which has been used in a number of other studies (Gross and Levenson, 1995; Goldin et al., 2005; Hutcherson et al., 2005; Britton et al., 2006), and from ‘A Perfect World’ (Johnson, 1993; duration 11 min and 30 s). ‘A Perfect World’ was selected from a database of emotion-eliciting film clips from the Université Catholique de Louvain (Schaefer et al., 2010) composed of 64 clips of frequently mentioned scenes ranked on several affective dimensions. For both movies, we selected clips that contained moments of happiness and moments of negative affect to ensure that our time-series would contain significant variance in affect. The segments extracted from ‘A Perfect World’ contained both positive and negative emotions (ranked 48th and 54th percentile in the database, respectively), and scored highly on arousal (92nd percentile). For the Champ, we had

no such validated ratings of each segment, but we included segments of similarly intense positive and negative affect. Both clips were expected to induce synchronization of emotional brain regions across participants because of their high arousal and clearly discernible sequence of positive and negative emotions they elicit in viewers. The video clips were presented in a counterbalanced order using Presentation software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, [www.neurobs.com](http://www.neurobs.com)) running on a Windows platform. They were front-projected via an LCD projector onto a screen placed at the back of the scanner bed, observable to participants via a mirror mounted within the head coil. The soundtrack was audible via an fMRI compatible headphone set.

### 2.4. Behavioural data processing

Ratings were acquired at 50 Hz, and downsampled at 0.5Hz by averaging the coordinates within each 2s time bin, to match the temporal resolution of the fMRI data.

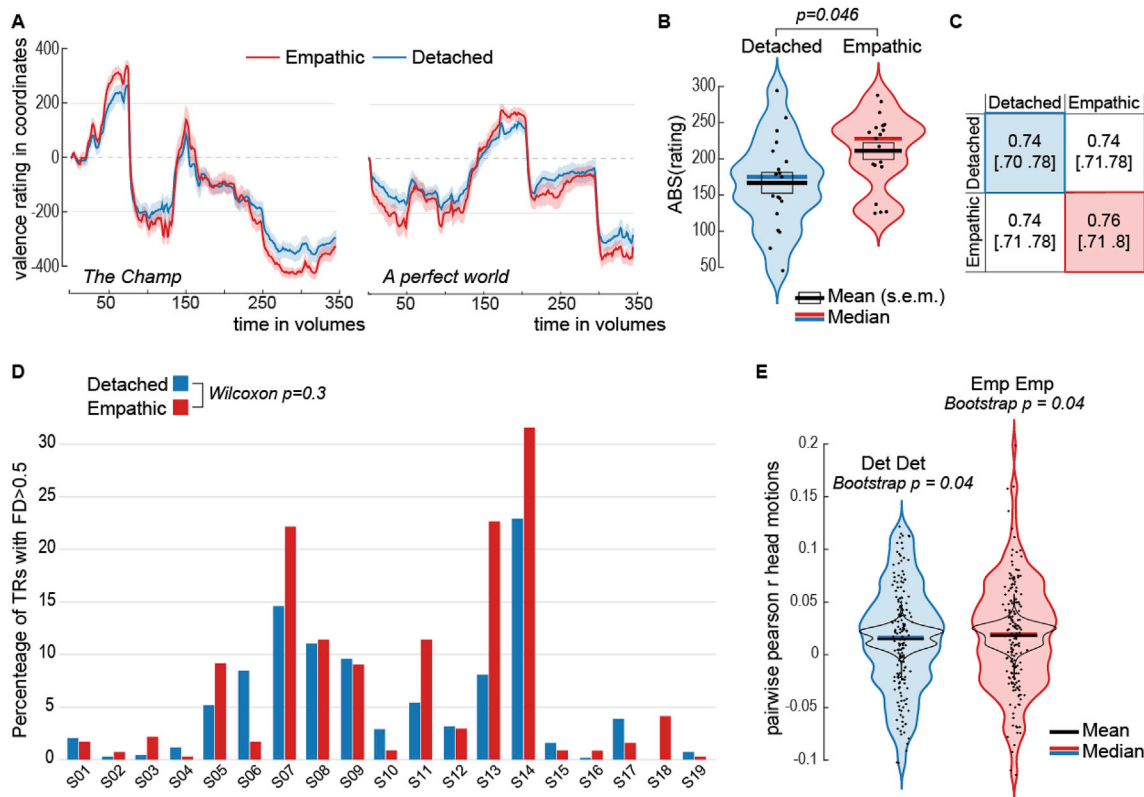
For each participant and video clip separately, we first computed simple correlation values between the ratings under each instruction - i.e.  $r(\text{rating}_{\text{Detached}_i\text{Clip1}}, \text{rating}_{\text{Empathic}_i\text{Clip1}})$  and  $r(\text{rating}_{\text{Detached}_i\text{Clip2}}, \text{rating}_{\text{Empathic}_i\text{Clip2}})$ . As our interest was to investigate the effects of our task instruction independently of a specific video, rather than differences between video clips, the correlation values were then z-transformed and averaged across video clips to test the significance of the correlation over zero with simple T-tests. To identify differences across instructions, we also calculated the mean rating (over time and video clip), the mean absolute value of the rating and its standard deviation (over time), and these values were compared between instructions using paired T-Tests, after verification of normality (all Shapiro tests for normality had  $p > 0.05$ ).

To quantify how consistent the ratings were across participants, we then ran a pair-wise Inter-Subject Correlation (ISC) (i.e.  $r(\text{rating}_{S_i}, \text{rating}_{S_j \neq i})$ , with  $S = \text{subject } i \text{ to } j$ ) for each video clip and task instruction, separately. The correlation values were then z-transformed and again averaged over video clips to obtain a cross-correlation matrix per task instruction. Because the consistency (r-value) was similarly high under both instructions, we then asked ourselves how similar the actual ratings were across conditions. To do so, we plotted the average ratings for each condition (Fig. 1A), and calculated additional pairwise ISC values across instruction, i.e. by correlating the individual ratings during the Detached instruction with other individuals’ rating during the Empathic instruction, and vice versa, by correlating the individual ratings during the Empathic instruction with those during the Detached instruction. Once again, the correlation values from the two video clips were z-transformed before averaging them. To assess the significance we perform a non-parametric test on the pairwise behavioral ISC. We do this with a bootstrap procedure, as especially recommended for this type of data (Chen et al., 2016). This procedure provides us with a p-value and with a confidence interval for the chosen statistical estimator, which is the median ISC.

### 2.5. MRI data acquisition

Two functional and one structural MR scans were acquired per session at the Spinoza Center at the University of Amsterdam, using a 3-T Philips Achieva 3.0 scanner equipped with a 32-channel head-coil for approximately 45 min total, with short pauses in between acquisitions in which participants could relax while laying still inside the scanner. Subjects were provided with a microphone and headphones to maintain communication throughout the scanning sessions. Head cushions inside the coil were positioned around the subjects’ head to minimize motion.

In each scan session, an anatomical scan was acquired using T<sub>1</sub>-weighted images (220 slices; TR = 8.2 ms; TE = 3.8 ms; inversion time 670.4 ms; FOV = 240 × 188 mm<sup>2</sup>, matrix size = 240 × 240; flip angle = 8°; voxel size = 1 mm<sup>3</sup>). Whole-brain functional T<sub>2</sub>\*-weight MRI data were acquired using a single-shot, ascending, gradient echo, echo planar,



**Fig. 1. Behavioral results.** (A) Average rating ( $\pm$ sem, in bold and light lines, respectively) across participants as a function of time, separately for the two video clips. The x-axis is in volumes of MRI acquisition and each volume corresponds to 2 s, emotional valence ratings are expressed in the  $-500$  to  $500$  range of the visual analogue scale. (B) Violin plots of the absolute value of the ratings across participants as a measure of the rating excursion. Each dot represents the average absolute value for one participant. (C) Pairwise ISC r-values across conditions. Each cell indicates the median r-values and the 95% confidence interval. The values on the diagonal represent the r-values within a condition, while the values off diagonal represent the ISC across conditions. (D) Percentage of volumes with frame-wise head displacements over  $0.5$  mm for each participant (S01–S19) and task instruction. (E) Violin plot of the pairwise correlations of the frame-wise head displacements for the Detached and Empathic instruction. The dots are the r values of all computed pairwise correlations ( $19 \times 19$ ), the black violin contours are the distributions of the boot-strap procedure used to calculate the significance of the observed average correlations. Data can be found at [https://www.dropbox.com/s/7y23h6adljnvdnq/Ratings\\_%26\\_FrameWiseHeadDisplacements.xlsx?dl=0](https://www.dropbox.com/s/7y23h6adljnvdnq/Ratings_%26_FrameWiseHeadDisplacements.xlsx?dl=0).

imaging sequence (voxel size =  $3 \text{ mm}^3$ ; TR = 2000 ms, TE = 29.93 ms; flip angle =  $76.1^\circ$ ; FOV =  $240 \text{ mm}^2$ ). Thirty-seven, 3-mm thick, transverse slices, with a 0.3 mm gap per run were acquired, allowing a full brain coverage in all subjects. Three-hundred and seventy volumes were collected per clip, resulting in 740 vol per session. Ten seconds of gradient RF pulses preceded the acquisition of each of the experimental runs to establish steady-state tissue magnetization, these images were automatically discarded from the raw data and were not included in any of the analysis.

## 2.6. MRI data preprocessing

Brain images were reconstructed from k-space and par/rec format files were converted to nifti format and subsequently preprocessed in SPM12 software (Wellcome Trust Centre for Neuroimaging, University College London, UK <http://www.fil.ion.ucl.ac.uk/spm/>) run with Matlab R14b ([www.mathworks.com](http://www.mathworks.com)) on a Windows platform. The first five functional images of each run, corresponding to the instructions slides, were removed to include only those volumes where participants were observing and rating the clips, resulting in 361 and 344 vol for 'The Champ' and 'A Perfect World', respectively. Firstly, to adjust between-slice time differences, functional images were phase-shifted using the slice-time correction utility in SPM to the timing of the middle slice of each volume using a sinc interpolation method, and then spatially realigned to the first image acquired using an iterative rigid body transformation and a least square approach to correct for movement artefacts. Realignment parameters were used to (i) calculate the percentage of

framewise displacement (FD) above  $0.5$  mm (Fig. 1D) for each participant and condition, and (ii) compute a pairwise ISC of the FD time series to explore whether head movements were synchronized across subjects (Fig. 1E). As significant correlations were identified by bootstrapping with 1000 possible permutations, we then regressed out the FD in the ISC and ISFC analyses (see below). The anatomical image was then co-registered to the mean EPI and segmented. The forward normalization parameters estimated during segmentation were saved and applied to both the anatomical and functional images for the spatial normalization to the T1-weighted Montreal Neurological Image space (MNI template). Functional and anatomical images were resampled into  $(2 \text{ mm})^3$  and  $1 \text{ mm}^3$  voxels respectively, and the default settings of SPM12's bounding box were adjusted [ $-90 -126 -72; 90 90 108$ ] in order to avoid omission of cerebellar voxels (Gazzola and Keysers, 2009; Abdelgabar et al., 2018). A 6 mm full-width at half maximum (FWHM) Gaussian kernel was applied to the functional images to reduce inter subject anatomical variability.

## 2.7. fMRI intersubject correlation (ISC)

We compute pairwise ISC within and across instruction (Empathic and Detached), and test statistical significance through non-parametric testing. We do this for the data corresponding to each clip independently. Subsequently, we averaged the pairwise correlation matrices of the two clips as a way to filter out clip-specific information. For each voxel and combination of instruction (e.g. Empathic, Empathic), we therefore have a  $N_{\text{sub}} \times N_{\text{sub}}$  matrix ISC(Empathic, Empathic), containing the correlations of the timecourses measured at that voxel for each pair of

subject ( $S_i, S_j$ ). As an estimator for the correlation at each voxel and between instruction we compute the median of the upper triangular part of this symmetric matrix and assess its significance through non-parametric testing. Note that we only focus on ISCs and exclude the diagonal, which is trivial for the within-condition case. The non-parametric test that we employ is the one suggested in Chen et al. (2016) and publicly available in the ISC module of the brainiak project (<https://brainiak.org/docs/brainiak.html#module-brainiak.isc>). Specifically, we use the bootstrap\_isc method. This method generates a bootstrap distribution for the median  $r$  value of the correlation matrix  $M(\text{emp}, \text{det})$  by randomly replacing the rows and columns corresponding to one or more subjects, with those of other subjects. This operation can then be repeated an adequately large number of times to generate a distribution of values that represents the variations of the original  $r$  and allows the estimation of a p-value. For more details please refer to the official documentation at [https://brainiak.org/docs/brainiak.html#module-brainiak.isc.bootstrap\\_isc](https://brainiak.org/docs/brainiak.html#module-brainiak.isc.bootstrap_isc). In our case, we performed  $N = 1000$  bootstrap. This gives us a median  $r$ -value per voxel, each with its corresponding p-value. We repeat all of this for the four possible cases in which we evaluate the significance of the ISC in one pair of instructions against the null hypothesis (Fig. 2), combining the information corresponding to the two clips as described above. Then, we perform the same operation for the contrasts between

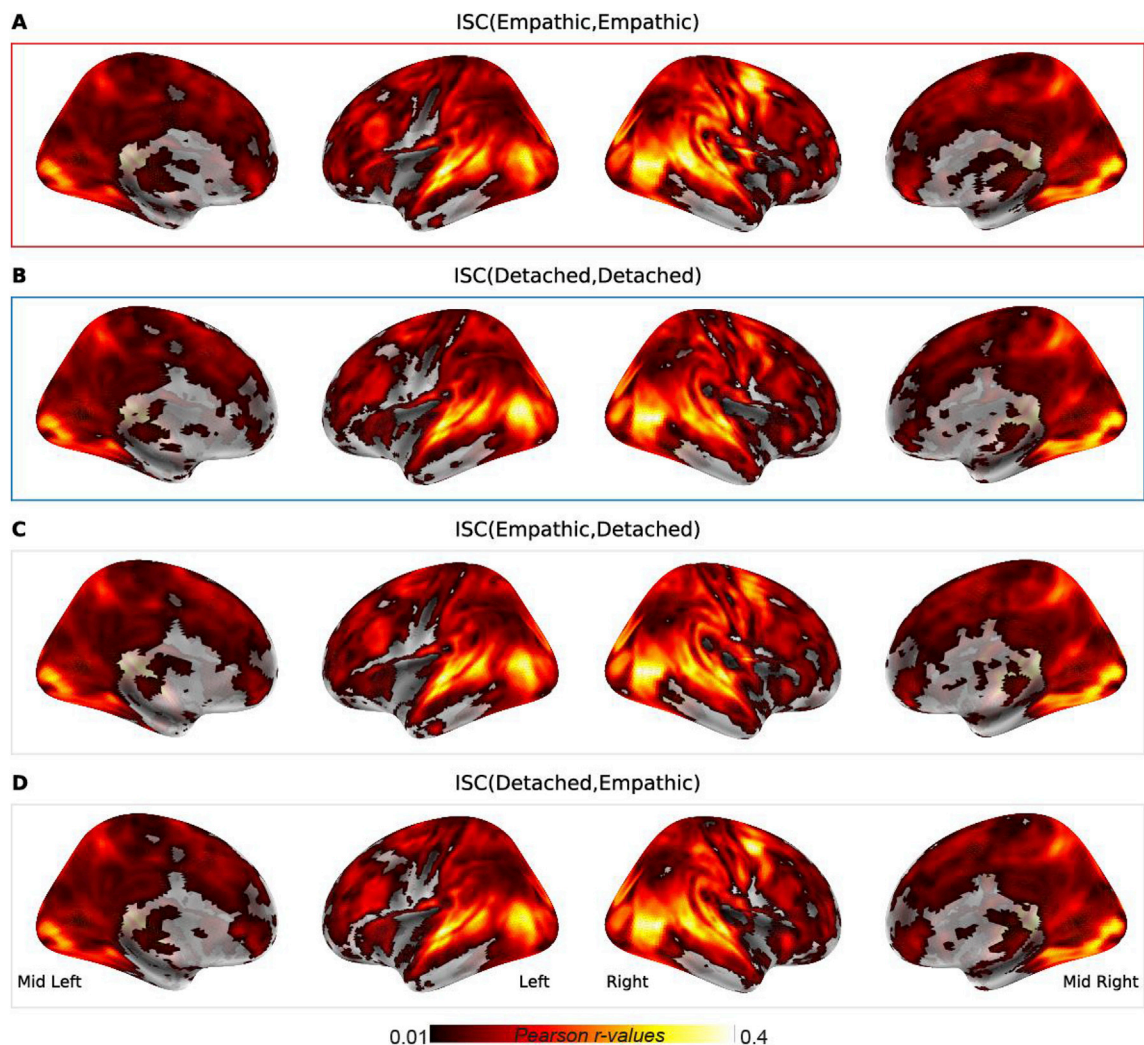
different instructions (Fig. 3), where we combine the correlation matrices corresponding to different clips and instruction, weighting them with a  $\pm$  sign, according to the desired contrast. For example, in the case of Empathic > Detached we have:

$$r(\text{Emp} > \text{Det}) = 1/4[r(\text{clip1}; \text{Emp}, \text{Emp}) + r(\text{clip2}; \text{Emp}, \text{Emp}) - r(\text{clip1}; \text{Det}, \text{Det}) - r(\text{clip2}; \text{Det}, \text{Det})]$$

To take into account the significant ISC found for head movements (i.e. FD), we performed the ISC calculation excluding possible additional correlations caused by such movements. We did this by fitting a general linear model for the BOLD response as a function of FD, implementing the fit for  $\text{FD}_{(t)}$ ,  $\text{FD}_{(t-1)}$ ,  $\text{FD}_{(t+1)}$ , and their squared values. We then retained the residuals of this fit as the “true” signal, to correlate across every possible pair of subjects and instruction for our ISC study. Including this extra step in the ISC analysis reassuringly does not lead to significant differences neither in the ISC corresponding to a single pair of subjects, nor in the groupwise analysis presented in Figs. 2 and 3.

## 2.8. fMRI intersubject functional connectivity (ISFC)

To explore whether the instruction to deliberately empathize or detach from the main character’s feelings temporally synchronized the connectivity of networks across participants, we implemented an



**Fig. 2. Pairwise ISC of brain activity.** Spatial representation of the synchronization of BOLD activity (in terms of intersubject pair-wise correlation) for the Empathic and Detached instructions. Activations are rendered with the plot\_surf package of Nilearn ([https://nilearn.github.io/modules/generated/nilearn.plotting.plot\\_surf.html](https://nilearn.github.io/modules/generated/nilearn.plotting.plot_surf.html)), sampling the data on a cortical mesh from the standard Nilearn datasets (inflated brain from [https://nilearn.github.io/modules/generated/nilearn.datasets.fetch\\_surf\\_fsaverage.html#nilearn.datasets.fetch\\_surf\\_fsaverage](https://nilearn.github.io/modules/generated/nilearn.datasets.fetch_surf_fsaverage.html#nilearn.datasets.fetch_surf_fsaverage)). Only voxels surviving a bootstrap non-parametric test at  $q < 0.01$  FDR corrected are shown. All data shown are available at this link <https://identifiers.org/neurovault.collection:6079>.

adaptation of the ISFC analysis (Simony et al., 2016). The ISFC explores how information about stimuli (Nastase et al., 2019) is exchanged across brain networks by exploring the temporal correlation between the time course of a network in one participant, and the time course of another participant in a different network. Doing this analysis across participants ensures that correlations are not due to confounds within a particular brain (e.g. heart beat, respiration, or other intrinsic signals), but can only originate from information about the movie (Nastase et al., 2019; Simony et al., 2016). For this, the EPI volumes of each clip and condition were band-pass filtered (0.1–0.01 Hz) to isolate the temporal variations between 10 and 100 s which are expected to encode most of the BOLD signal relevant for cognitive processes (Honey et al., 2012), standardized, averaged across all participants, to generate one time series per movie and condition, and then temporally concatenated into a file containing both clips and conditions. Then, we used an ICA (FSL, Melodic algorithm from Beckmann and Smith, 2004; Smith et al., 2004) to decompose brain activity into 20 large scale networks (Fig. 4) that had representations of the movie that are consistent across participants. The ISFC was computed across each pair of independent components (ICs) and for every possible pair of participants. Additionally, we made sure to exclude any possible spurious correlations caused by head motion artefacts by computing ISFC as a partial correlation  $r(IC_a S_i, IC_b S_j | FDS_i, FDS_j)$ , where  $IC_a$  and  $IC_b$  are two possible IC pairs,  $FDS_i$  and  $FDS_j$  the head displacement of subject  $i$  and subject  $j$ . Note that for FD we include also the same time series shifted back and forward by one TR, and their squared values, both for subject  $i$  and  $j$ , for a total of 12 confound variables. With this procedure, we compute  $N_{sub} \wedge 2$  connectivity matrices, one for each pair of participants, from which we compute an average connectivity map, averaging across all of these pairs. Again, assessing the significance of the measured values is a non-trivial problem, given that we are using a pairwise approach (Chen et al., 2016). Ultimately, we were interested in the difference in connectivity between the two studied conditions, empathic and detached. To assess the significance of this difference, we implemented a pairwise permutation test in which, for each subject, we flipped the Empathic vs Detached label, and then recalculate the entire pairwise ISFC matrix of correlations between each  $IC_a$  and  $IC_b$ , and then compute the difference in mean pairwise correlation for Empathic-Detached. We did that 5000 times to get a distribution of random differences in correlation. To deal with the multiple comparison problem across the 190 combinations of two ICs, we then looked for a critical difference in

r-value, 'rho', so that in less than 5% of permutations we have any of the 190 pairs of IC pairs that change their correlation by 'rho'. At this critical rho value, FDR is thus 0.05.

### 2.9. Data and code availability statement

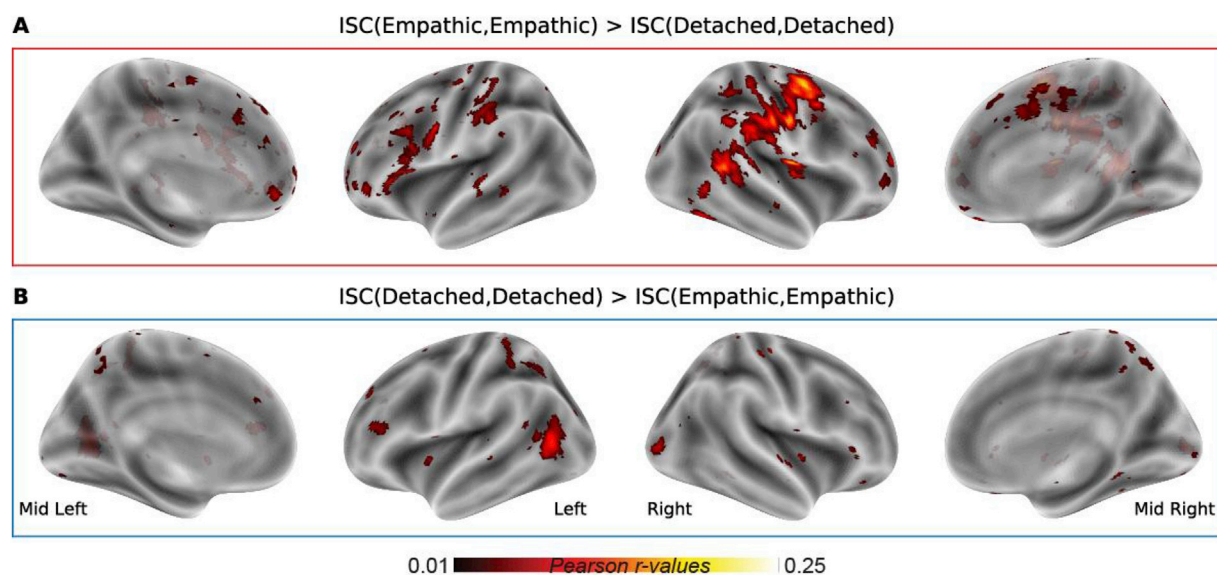
Raw data is available upon direct request to the corresponding author. The summary statistics that are shown in Figures can be downloaded at: <https://www.dropbox.com/sh/y609y0a484m1nr1/AADLCYurIbk-Kko81KwpwIR4a?dl=0>.

The data and code sharing adopted by the authors comply with the requirements of our funding bodies, the Netherlands Institute of Neuroscience, and comply with the Ethics Review Board Committee of the University of Amsterdam.

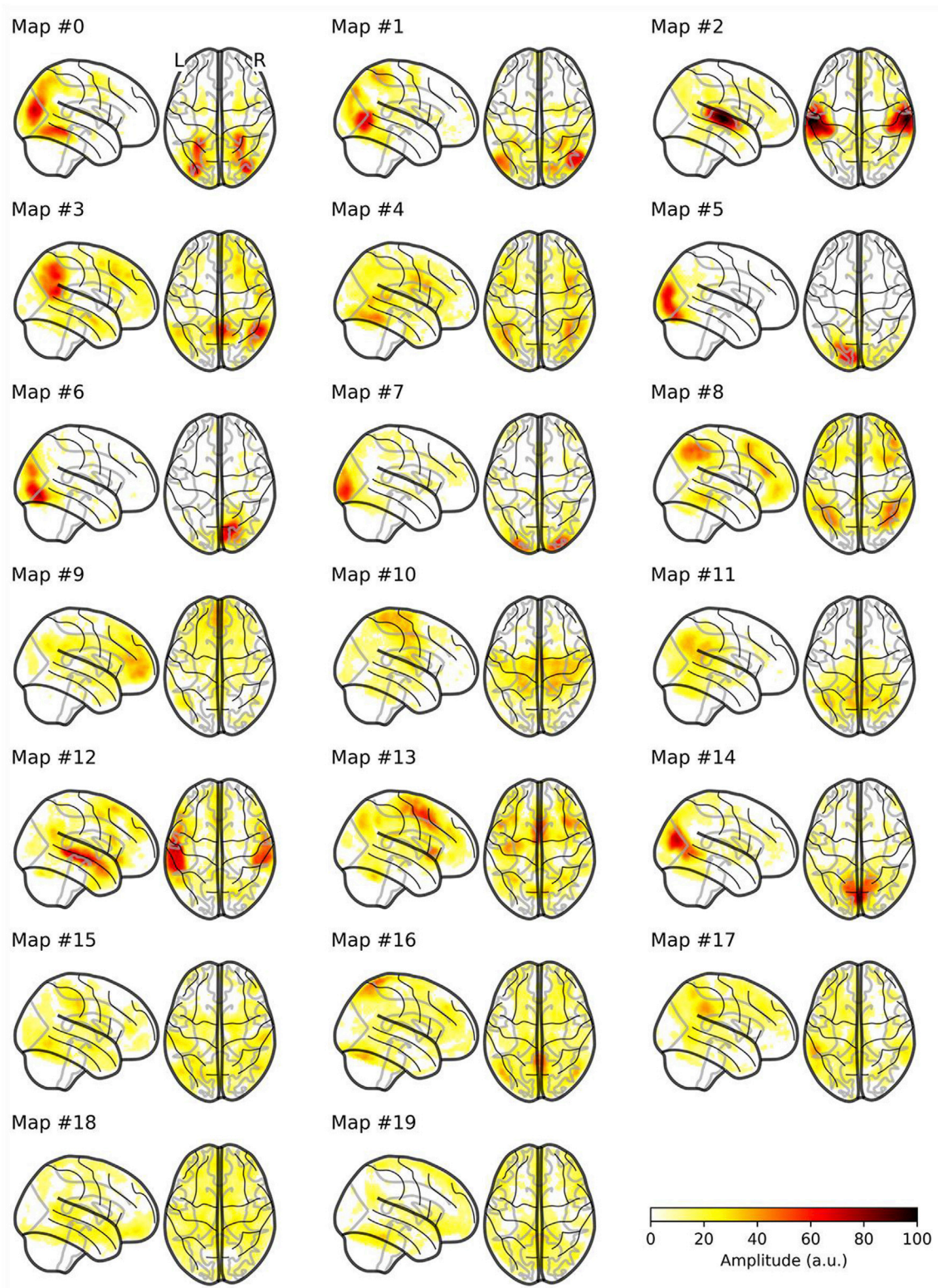
## 3. Results

### 3.1. Behaviour

To explore the impact of instructions on the participant ratings of the main character's feelings, we firstly calculated a simple correlation between the Empathic and the Detached ratings at the participant level, averaging over the two video clips. The average correlation was high (mean  $r = 0.85$ ,  $SD = 0.02$ ). A T-test confirmed that the correlation strongly differed from zero ( $t_{18} = 21.33$ ,  $p < 0.001$ ). Visually inspecting the average rating under the Detached and the Empathic instructions (blue and red curves, Fig. 1A) illustrates how much the two curves go up-and-down together. However, the Detached curve appears to be flatter than the Empathic curve. To capture this difference in excursion, we explored whether there were differences in the mean rating, and in the excursion of the rating. The mean ratings between conditions differed slightly (mean rating Detached =  $-112.49 \pm 60.02$  SD; Empathic =  $-133.63 \pm 47.26$  SD;  $t_{18} = 2.1$ ,  $p = 0.046$ ). The excursion differed more substantially, which was evident when quantifying the mean absolute value of the rating (Fig. 1B; average absolute rating Detached =  $167.7 \pm 63.8$  SD, Empathic =  $212.2 \pm 52.12$  SD;  $t_{18} = -4.15$ ,  $p < 0.001$ ), which captures how far from the neutral state participant rated the character on average (independently of direction). This was also evident when comparing the standard deviation over time of the ratings of each participant (average standard deviation Detached =  $158.8 \pm 50.04$  SD,



**Fig. 3. Instruction-dependent ISC.** (A) Regions where ISC is larger under Empathic than under Detached instructions. (B) Regions where ISC is larger under Detached than under Empathic instructions. Activations are rendered using the Nilearn library, as for Fig. 2. The similarity between the contrast maps and functional networks associated with embodied cognition and mentalizing is described in Fig. S4 and Fig. S5. All data shown are available at this link <https://identifiers.org/neurovault.collection:6079>.



**Fig. 4. ICA maps.** Glass brain representation of the 20 ICs (Map 0 to Map 19) obtained using Melodic on the average concatenated standardized time courses of the two video clips under the two instructions. Please note that the numbering of the ICs is the same as in Fig. 5. The similarity between each IC and functional networks associated with embodied cognition and mentalizing is described in Fig. S4 and Fig. S6. All data shown are available at this link <https://identifiers.org/neurovault.collection:6079>.

Empathic =  $200.8 \pm 42.25$ ;  $t_{18} = -4.23$   $p < 0.001$ ). Taken together, these results indicate that although participants capture the emotional up-and-downs of the protagonist similarly under the two conditions, under the Empathic condition participants yielded more extreme ratings

than they did under the Detached condition.

Fig. 1C shows the results of the pairwise ISC analysis that was ran on the behavioural data. Results show a high agreement across participants under both the Detached and Empathic conditions. The ISC(Detached,



Detached) values fall within the confidence intervals of the ISC(Empathic, Empathic) and this shows that emotion ratings were not less reliable under the Detached condition, as would have been expected if our participants paid less attention to the emotions of the protagonist. The same is true when restricting the analyses to those participants that performed under a given instruction in the first session (ISC(Detached, Detached) = 0.74 [0.63 0.84], ISC(Empathic, Empathic) = 0.70 [0.56 0.86]), showing that the similarity in ISC values is not due to having already rated the movie under the other instruction beforehand. Because ratings could be consistent within each instruction, but different across instructions, we also computed ISC across conditions: ISC(Detached, Empathic) and ISC(Empathic, Detached). Again, as the confidence interval overlaps we can conclude that across participants the ratings across conditions were as similar as those within conditions.

Fig. 1D shows the percentage of framewise head displacement present in our data separately for the two sessions. As the Shapiro-Wilk test for normality on these percentages was close to significance ( $p = 0.054$ ), we used a non parametric Wilcoxon test to compare conditions. The test did not reveal significant difference across instructions ( $p = 0.3$ ). A pairwise ISC on the time-series of framewise displacements revealed that displacement was weakly but significantly correlated across individuals in both the Detached and Empathic conditions (Fig. 1E). We therefore regressed out signals correlated in time with the displacement in all the following brain analyses (see method section 2.7 and 2.8).

### 3.2. Emotional stimuli synchronize brain activity

The pairwise ISC that was computed for the Empathic and Detached instructions show that movies reliably synchronized a broad network of brain regions that included, but were not limited to, early auditory and visual regions as well as somatosensory, parietal, premotor, prefrontal, and limbic cortices (Fig. 2A and B). Fig. 2C and D additionally show the circuit consistently recruited during cross-condition (i.e. ISC(Detached, Empathic) or ISC(Empathic, Detached)) is very similar to the within condition ISC maps (i.e. ISC(Detached, Detached) or ISC(Empathic, Empathic)) confirming that, in line with what we observe in the ratings, the time course of brain activity is similar across conditions in several brain regions (for more details on these regions please refer here).

### 3.3. Instructions change the ISC in the brain

To test if instructions alter the synchronization of brain regions we compared the ISC values between conditions (Fig. 3). Results confirmed that instructions altered the brain regions used to encode the movies. Regions that synchronize more in the Empathic compared to the Detached condition include the inferior, middle, and superior frontal gyri, the premotor cortex, the cingulate cortex, as well as the primary and secondary somatosensory cortices, the inferior parietal and the right posterior inferior temporal cortices (Fig. 3A and [Inline Supplementary Table 1](#) for a more detailed list of activations). Clusters that synchronize more during the Detached compared to the Empathic condition, include but are not limited to occipito-temporal, superior parietal, and inferior frontal regions (Fig. 3B and [Inline Supplementary Table 2](#) for a more comprehensive list of activations).

### 3.4. Instructions change the functional connectivity within the brain

To explore whether the instructions changed functional connectivity across the large scale networks representing the movies, we implemented an adaptation of intersubject functional connectivity (ISFC; [Nastase et al., 2019](#); [Simony et al., 2016](#)).

We used an ICA to decompose and summarize brain activity into 20 large scale networks (Fig. 4) that have representations of the movie that are consistent across participants. We manually set the number of ICA to 20 networks, as has often been done for rsfMRI ([Smith et al., 2009](#)), to strike a balance between explaining sufficient variance and generating a

small but reproducible set of large-scale networks to maintain sensitivity by limiting the number of comparisons we need to correct for.

Subsequently, we computed the Pearson-correlations between each pair of the 20 identified ICs components and each pair of participants, partialling out possible spurious correlations due to FD. We compared these ISFC values across the two conditions and determined the significance of the ISFC variation with a non parametric test based on random flips of the Empathic vs Detached label ( $q = 0.05$ , see methods section 2.8). The ISFC analysis reveals that instructions significantly reconfigured functional connectivity across large scale brain networks (Fig. 5). A closer examination of the changes shows that in 83% of the significant changes, the instruction to empathize increases the ISFC across the pair of ICAs (i.e.  $|\text{ISFC}(\text{Empathic})| > |\text{ISFC}(\text{Detached})|$ ). The most frequent change is a positive ISFC(Detached) becoming even more positive under the Empathic instruction (warm colors in all the panels A–C of Fig. 5; 53% of all significant changes). Changes in the sign of the ISFC across conditions were rare (warm colors in Fig. 5A, and cold colors in Fig. 5B or vice versa; 19% of all changes). Instructions to Empathize thus lead to a strengthening of functional connectivity across large scale networks while attributing emotions to others.

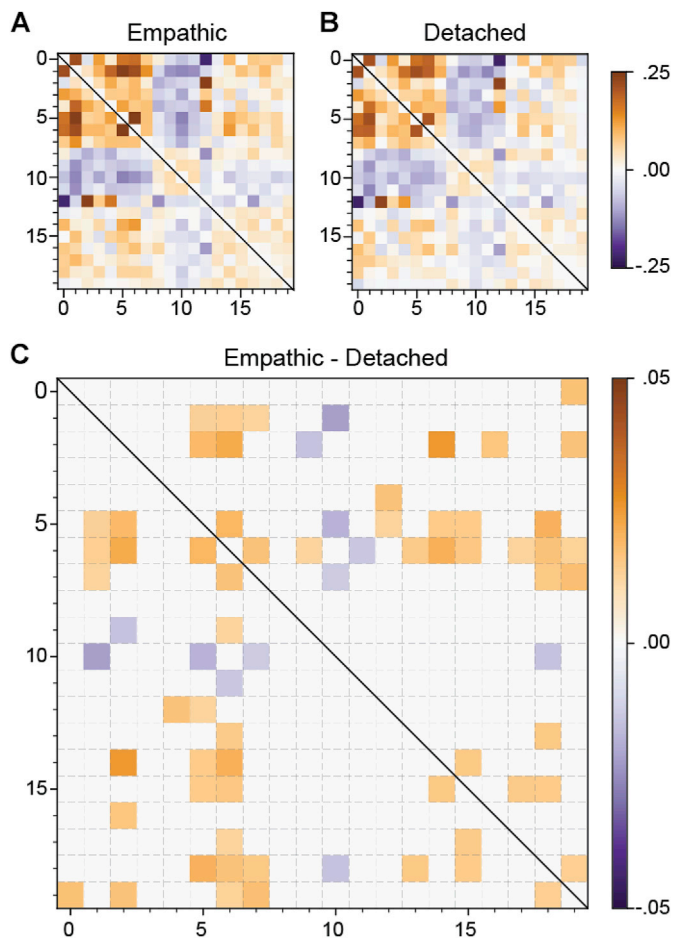
## 4. Discussion

To investigate whether and how participants have voluntary control over the degree to which they empathize with the emotions of others, we measured brain activity while participants watched emotional Hollywood movies under two different instructions: to rate the main characters' emotions by empathizing with them, or to do so while keeping a detached perspective. We found that participants yielded highly consistent and similar ratings of emotions under both conditions. Using intersubject correlation-based analyses we found that, when encouraged to empathize, participants' brain activity in limbic (including cingulate and putamen) and somatomotor regions (including premotor, SI and SII) synchronized more during the movie than when encouraged to detach. Using intersubject functional connectivity we found that comparing the empathic and detached perspectives revealed widespread increases in functional connectivity between large scale networks.

Participants were highly consistent in their rating of the protagonists' feelings: this was true when comparing the ratings across participants and across the two instruction conditions ( $r > 0.74$ ). Such high consistency is perhaps not surprising, considering that the movies were chosen to depict strong and clear-cut emotions. That the consistency of ratings across participants was indistinguishable across the two instructions is important, as it suggests that participants were similarly attentive to the feelings of the character under both conditions: had they been more distracted under the Detached condition, we would have expected an increase in noise, hence a reduction of consistency. However, we also find that instructions to empathize amplified the excursions in the ratings: while participants agreed across conditions about when the character was feeling more positive or more negative (as quantified using correlations, which are scaling invariant), they felt the 'ups' were more positive and the 'downs' more negative under the instructions to empathize. This shows that the voluntary modulation of empathy has an impact on the self-reported perception of other people's feelings.

With regard to brain activity, we found that the movies synchronized a broad and similar network of brain regions under both types of instructions. Most relevant to the question of whether participants are able to deliberately modulate empathy, however, are the differences we observed in the brain activity synchronization across instructions.

We had hypothesized that instructions to empathize should increase synchrony in somatosensory, motor and limbic structures associated with embodied processes. This was indeed the case: we found that instructions to empathize enhanced synchronization in the primary and secondary somatosensory cortices, inferior parietal, inferior frontal and premotor cortices. All of these regions have been shown to be active both while participants perform actions and observe them ([Gazzola et al., 2007a](#),



**Fig. 5. Instructions alter Intersubject Functional Connectivity.** Connectivity maps for the 20 ICs in the Empathic condition (A) and Detached condition (B). In (C) the changes in connectivity (Empathic-Detached) - only the differences that are significant are shown (FDR  $q = 0.05$ ), the rest is set to 0. Data can be found at <https://www.dropbox.com/sh/ygqsdfrie5jc9q/AAB5KQVAMYjp6ntiKeTCwUHa?dl=0>.

2007b, 2009; Rizzolatti and Craighero, 2004), or while experiencing and observing touch (Keysers et al., 2004, 2010; Keysers and Gazzola, 2009). The homologous brain regions in monkeys have been shown to contain mirror neurons that map the actions and sensations of others onto the monkey's own (Rizzolatti and Sinigaglia, 2010; Gallese et al., 2004; Umiltà et al., 2001); and deactivating these brain regions impairs the ability to perceive subtle kinematic cues (Valchev et al., 2017; Pobric and Hamilton, 2006). This is compatible with the notion that while participants judge the actions and emotions of others, they can deliberately regulate the degree to which they allow their own actions and sensations to resonate with those they observe. In addition, we found that a number of limbic brain regions also increase synchrony under the instructions to empathize. This included in particular the mid-cingulate cortex and the putamen. In the cingulate, this increased synchronization falls within the midcingulate regions associated with the first hand experience of pain and with witnessing the pain of others (Engen and Singer, 2013; Keysers et al., 2010; Lamm et al., 2011; Singer et al., 2004; Meffert et al., 2013), and mirror neurons for pain have recently been reported in the homologous region of the rat (Carrillo et al., 2019). This is compatible with the notion that participants can voluntarily determine how much they allow the pain of others to influence representations of their own suffering. Activity in the dorsal striatum, including the putamen, has been associated with reinforcement learning, and has been shown to be recruited both while experiencing success and witnessing other people's success (Monfardini et al., 2013). This activity could therefore be compatible

with the notion that participants under the Empathic condition let the reward signals experienced by the protagonists recruit their own reinforcement learning systems more than during the detached condition. Jointly, these observations are in line with the psychological literature showing that empathy is a deliberately modulated process (Keysers and Gazzola, 2014; Zaki, 2014; Weisz and Zaki, 2018).

We had also hypothesized that instructions to detach might increase synchronization in regions associated with mentalizing, such as the TPJ, pre-cuneus and vmPFC (Mar, 2011; Saxe and Kanwisher, 2003; Schurz et al., 2014). Comparing our Detached-Empathic results with the meta-analysis of Mar, 2011 indeed reveals that regions with increased synchronization under instructions to detach overlapped with the mentalizing network identified, particularly for non story-based tasks, including in the TPJ and pre-cuneus (Fig. S5). In contrast, the Emp > Det contrast shared more similarity with activations associated with networks associated with embodied concepts (Fig. S5). This trade-off in ISC between regions associated with mentalizing and regions involved in embodied processes confirm the proposal that instructions can alter the balance across embodied and cognitive routes to the emotions of others (Keysers and Gazzola, 2007).

That we used ISC based methods to examine the voluntary control of empathy has implications for the interpretation of the results. Had we made a block design with short empathy-triggering stimuli asking participants to either empathize or detach on individual blocks, the contrast of blocks in the Empathize and Detach conditions would have identified both regions that control empathy (and would be tonically active during blocks of a particular condition) and regions that represent the emotions of the characters. Using ISC, we do not compare the average activity across conditions (which are removed during demeaning), but focus on fluctuations in time within a clip that are time-locked to the clip and hence carry information about the content of the movie (see Nastase et al., 2019 for a detailed discussion of what ISC measures). In this particular context, this means that differences in ISC across conditions will primarily identify differences in how the movie is represented rather than tonic task processes. This makes our results complement those from more traditional designs, including those that ask people to regulate their own emotions while viewing IAPS pictures that are often social in nature (see Kohn et al., 2014; Ochsner et al., 2012 for reviews).

In addition to investigating ISC, we also measured the impact of instructions on intersubject functional connectivity. Because investigating the connectivity between every possible pair of voxels would lead to an explosion of multiple comparisons, we decided to reduce the dimensionality of our dataset by first using an ICA on the voxels' time course averaged over participants. Averaging over participants before the ICA ensured that the ICA preserved as much information as possible about the activity fluctuations triggered by and hence time-locked to the movies. This analysis revealed that the flow of stimulus-relevant information was modified by the instructions, with the most frequent effect being that instructions to empathize increased the functional connectivity across networks relative to instructions to detach. Because of the number and extent of the networks involved, a detailed discussion of the function of each network that takes part in increased connectivity would seem excessive. Comparing the ICs with networks associated with embodied and mentalizing networks (Fig. S6) shows that some of the networks that load most strongly on embodied networks (e.g. ICs 1, 10 and 15), change their connectivity substantially depending on instructions to empathize, while networks that load most strongly on mentalizing networks (e.g. ICs 3 and 9) maintain relatively unaltered connectivity.

Overall, that instructions to empathize lead to (i) stronger intersubject synchrony (i.e. the contrast Empathic > Detached revealed more significant differences than the reverse contrast), (ii) stronger intersubject connectivity and (iii) more extreme emotion ratings suggests that the instructions to empathize may work in the brain as an overall opening of neural gates that allow the movie to influence brain activity more systematically and along stronger connections and triggers more intense representations of the stimulus. Conversely, instructions to be detached

serve to close such gates, forcing some information through alternate networks and potentially serving as a protection mechanism.

Our study has a number of limitations that should be considered. Firstly, in previous approaches (Nummenmaa et al., 2012; Viinikainen et al., 2010), ratings about the emotions of the characters were not collected in the scanner, but later during a second viewing of the movie. In contrast, our participants were asked to rate the emotions of the character while brain activity was measured. This was done to ensure that participants would have to process the character's emotions in all conditions, as confirmed by the similar level of consistency (in terms of ISC) of the ratings across empathic and detached conditions. However, our approach has the disadvantage of creating a more artificial situation, in which having to generate reports may have altered brain activity relative to what it would have been during a more implicit emotion processing during an unconstrained viewing. Repeating our experiment without online rating could reveal how much of our instruction effects would then be observed. Second, in our paradigm, we asked participants to report how the main protagonist felt, therefore asking them to provide a perceptual judgement. It is difficult to evaluate to what degree participants felt the reported emotions. Psychophysiological measures could in the future be analysed using a similar ISC approach to explore whether the degree to which participants physiological states align to the movie would also be under the deliberate control of our participants. Third, we measured changes of ISC across instructions using the average ISC over both movies. This approach does not allow us to pinpoint the moments in the movies in which a particular voxel shows significant ISC or significant changes in ISC. The tentative associations of brain regions showing differential ISC and somato-motor or nociceptive content then remains highly tentative, given all the limitations of reverse inference (Poldrack, 2006). In the future, time resolved ISC analyses, that pinpoint the moment in time where ISC becomes significant (and significantly different across conditions) could provide further insights into what aspect of the stimulus is being processed in a given region.

In summary, we provide evidence that even when participants rate other people's feelings they are able to voluntarily gate the access of the stimulus material to embodied and mentalizing networks and the intensity with which they perceive the feelings of others. This adds to the extensive psychological literature that has demonstrated that participants do regulate their empathy to maximize the benefits of empathy and minimize its costs (see Weisz and Zaki, 2018 and Zaki, 2014 for recent reviews of this literature), and to the emerging brain imaging literature suggesting that brain activity is significantly altered while participants choose whether to focus on the social and empathy triggering aspects of complex emotional stimuli (Meffert et al., 2013; Bruneau et al., 2013; Lahnakoski et al., 2014) and that brain activity and structure can be altered by meditative practices that focus on different forms of intersubjective sensitivity (Klimecki et al., 2014; Valk et al., 2017).

#### Author contribution

VG and CK conceived and supervised the study. VG, CK, and KCBJ acquired the funding. Paradigm implementation, piloting and data collection was carried out by ARA and LM. KCBJ, ARA, LDA, CK and VG analyzed the data. KCBJ, ARA, CK and VG wrote the manuscript with help from all authors. All authors worked on the revision. VG and LDA generated the artwork. KCBJ, LDA and VG are responsible for data curation.

#### Data availability

Raw data will be made available in Zenodo upon publication. Data of Figs. 2–4 have been uploaded on Neurovault (<https://identifiers.org/neurovault.collection:6079>). Data of Figs. 1 and 5 can be found at [https://www.dropbox.com/s/7y23h6adljnvdnq/Ratings\\_%26\\_FrameWis eHeadDisplacements.xlsx?dl=0](https://www.dropbox.com/s/7y23h6adljnvdnq/Ratings_%26_FrameWis eHeadDisplacements.xlsx?dl=0) and <https://www.dropbox.com/sh/ygsg sdfrie5jc9q/AAB5KQVAMYj6pcentiKeTCwUHa?dl=0>.

#### Funding

This work was supported by the Netherlands Organization for Scientific Research (VICI: 453-15-009 to C.K. and VIDI 452-14-015 to VG), the Mexican National Council for Science and Technology (CONACYT, Becas de posgrado al extranjero, ref. 365840/382981 to KCBJ).

#### Declaration of competing interest

None.

#### Acknowledgements

We thank Uri Hasson for advice on analyses, Stephan Schulz and Tinka Beemsterboer for help in optimising the scanning parameters for the current study, Leonardo Cerliani for helping with the ICA implementation and FSL Randomise, Rajeev Rajendran, Steven Voges and Alessandra Nostro for their comments on the manuscript, and Ysbrand van der Werf for kindly providing the fMRI compatible trackball for this experiment.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2020.116529>.

#### References

- Abdelgabar, A.R., Suttrup, J., Broersen, R., Bhandari, R., Picard, S., Keyesers, C., De Zeeuw, C.I., Gazzola, V., 2018. Action perception recruits the cerebellum and is impaired in spinocerebellar ataxia patients. *Brain* (in press).
- Azevedo, R.T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., Aglioti, S.M., 2013. Their pain is not our pain: brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. *Hum. Brain Mapp.* 34, 3168–3181. <https://doi.org/10.1002/hbm.22133>.
- Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging* 23, 137–152. <https://doi.org/10.1109/TMI.2003.822821>.
- Bird, G., Silani, G., Brindley, R., White, S., Frith, U., Singer, T., 2010. Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain* 133, 1515–1525. <https://doi.org/10.1093/brain/awq060>.
- Britton, J.C., Phan, K.L., Taylor, S.F., Welsh, R.C., Berridge, K.C., Liberzon, I., 2006. Neural correlates of social and nonsocial emotions: an fMRI study. *Neuroimage* 31, 397–409. <https://doi.org/10.1016/j.neuroimage.2005.11.027>.
- Bruneau, E., Dufour, N., Saxe, R., 2013. How we know it hurts: item analysis of written narratives reveals distinct neural responses to others' physical pain and emotional suffering. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0063085> e63085.
- Cameron, C.D., Harris, L.T., Payne, B.K., 2016. The emotional cost of humanity. *Soc. Psychol. Personal. Sci.* 7, 105–112. <https://doi.org/10.1177/1948550615604453>.
- Carrillo, M., Han, Y., Migliorati, F., Liu, M., Gazzola, V., Keyesers, C., 2019. Emotional mirror neurons in the rat's anterior cingulate cortex. *Curr. Biol.* 29 (8), 1301–1312. <https://doi.org/10.1016/j.cub.2019.03.024>.
- Chen, G., Shin, Y.W., Taylor, P.A., Glen, D.R., Reynolds, R.C., Israel, R.B., Cox, R.W., 2016. Untangling the relatedness among correlations, part I: nonparametric approaches to inter-subject correlation analysis at the group level. *Neuroimage* 142, 248–259. <https://doi.org/10.1016/j.neuroimage.2016.05.023>.
- Cikara, M., Bruneau, E., Van Bavel, J.J., Saxe, R., 2014. Their pain gives us pleasure: how intergroup dynamics shape empathic failures and counter-empathic responses. *J. Exp. Soc. Psychol.* 55, 110–125. <https://doi.org/10.1016/j.jesp.2014.06.007>.
- Creemers, H.R., Wager, T.D., Yarkoni, T., 2017. The relation between statistical power and inference in fMRI. *PLoS One* 12 (11). <https://doi.org/10.1371/journal.pone.0184923> e0184923.
- Davis, M.H., 1980. *A Multidimensional Approach to Individual Differences in Empathy*. Dayan, E., Barliya, A., de Gelder, B., Hendler, T., Malach, R., Flash, T., 2018. Motion cues modulate responses to emotion in movies. *Sci. Rep.* 8 (1), 10881. <https://doi.org/10.1038/s41598-018-29111-4>.
- Decety, J., Meyer, M., 2008. From emotion resonance to empathic understanding: a social developmental neuroscience account. *Dev. Psychopathol.* 20 (4), 1053–1080. <https://doi.org/10.1017/S0954579408000503>.
- Engen, H.G., Singer, T., 2013. Empathy circuits. *Curr. Opin. Neurobiol.* 23, 275–282. <https://doi.org/10.1016/j.conb.2012.11.003>.
- Frith, C.D., Frith, U., 2008. Implicit and explicit processes in social cognition. *Neuron* 60 (3), 503–510. <https://doi.org/10.1016/j.neuron.2008.10.032>.
- Gallese, V., Keyesers, C., Rizzolatti, G., 2004. A unifying view of the basis of social cognition. *Trends Cogn. Sci.* 8 (9), 396–403. <https://doi.org/10.1016/j.tics.2004.07.002>.

- Gazzola, V., Rizzolatti, G., Wicker, B., Keysers, C., 2007a. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35 (4), 1674–1684. <https://doi.org/10.1016/j.neuroimage.2007.02.003>.
- Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., Keysers, C., 2007b. Aphasics born without hands mirror the goal of hand actions with their feet. *Curr. Biol.* 17 (14), 1235–1240. <https://doi.org/10.1016/j.cub.2007.06.045>.
- Gazzola, V., Keysers, C., 2009. The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebr. Cortex* 19, 1239–1255. <https://doi.org/10.1093/cercor/bhn181>.
- 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 (1), 26–36. <https://doi.org/10.1016/j.neuroimage.2005.03.018>.
- Gross, J.J., Levenson, R.W., 1995. Emotion elicitation using films. *Cognit. Emot.* 9, 87–108. <https://doi.org/10.1080/02699939508408966>.
- Gu, X., Han, S., 2007. Attention and reality constraints on the neural processes of empathy for pain. *Neuroimage* 36 (1), 256–267. <https://doi.org/10.1016/j.neuroimage.2007.02.025>.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. *Science* 303, 1634–1640. <https://doi.org/10.1126/science.1089506>.
- Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* 14 (1), 40–48. <https://doi.org/10.1016/j.tics.2009.10.011>.
- Hallam, G.P., Webb, T.L., Sheeran, P., et al., 2014. “The neural correlates of regulating another person’s emotions: an exploratory fMRI study.”. *Frontiers in Human Neuroscience* 8, 376. <https://doi.org/10.3389/fnhum.2014.00376>.
- Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrrod, S., Keysers, C., 2012. Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn. Sci.* 16 (114), 121. <https://doi.org/10.1016/j.tics.2011.12.007>.
- Hein, G., Silani, G., Preuschhoff, K., Batson, C.D., Singer, T., 2010. Neural responses to ingroup and outgroup members’ suffering predict individual differences in costly helping. *Neuron* 68, 149–160. <https://doi.org/10.1016/j.neuron.2010.09.003>.
- Honey, C.J., Theisen, T., Donner, T.H., Silbert, L.J., Carlson, C.E., Devinsky, O., Doyle, W.K., Rubin, N., Heeger, D.J., Hasson, U., 2012. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76, 423–434. <https://doi.org/10.1016/j.neuron.2012.08.011>.
- Hutcherson, C.A., Goldin, P.R., Ochsner, K.N., Gabrieli, J.D., Feldman Barrett, L., Gross, J.J., 2005. Attention and emotion: does rating emotion alter neural responses to amusing and sad films? *Neuroimage* 27, 656–668. <https://doi.org/10.1016/j.neuroimage.2005.04.028>.
- Jabbi, M., Swart, M., Keysers, C., 2007. Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* 34, 1744–1753. <https://doi.org/10.1016/j.neuroimage.2006.10.032>.
- Johnson, M., 1993. *Producer; Valdes D., Producer, Eastwood C., Director. A Perfect World [Film]*. Malpas Productions, United States.
- Keysers, C., Wicker, B., Gazzola, V., Anton, J.-L., Fogassi, L., Gallese, V., 2004. A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346. [https://doi.org/10.1016/S0896-6273\(04\)00156-4](https://doi.org/10.1016/S0896-6273(04)00156-4).
- Keysers, C., Gazzola, V., 2007. Integrating simulation and theory of mind: from self to social cognition. *Trends Cogn. Sci.* 11 (5), 194–196. <https://doi.org/10.1016/j.tics.2007.02.002>.
- Keysers, C., Gazzola, V., 2009. Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Curr. Opin. Neurobiol.* 19, 666–671. <https://doi.org/10.1016/j.conb.2009.10.006>.
- Keysers, C., Kaas, J.H., Gazzola, V., 2010. Somatosensation in social perception. *Nat. Rev. Neurosci.* 11, 417–428. <https://doi.org/10.1038/nrn2833>.
- Keysers, C., Gazzola, V., 2014. Dissociating the ability and propensity for empathy. *Trends Cogn. Sci.* 18, 163–166. <https://doi.org/10.1016/j.tics.2013.12.011>.
- Klimecki, O.M., Leiberg, S., Ricard, M., Singer, T., 2014. Differential pattern of functional brain plasticity after compassion and empathy training. *Soc. Cogn. Affect. Neurosci.* 9, 873–879. <https://doi.org/10.1093/scan/nst060>.
- Kohn, N., Eickhoff, S.B., Scheller, M., Laird, A.R., Fox, P.T., Habel, U., 2014. Neural network of cognitive emotion regulation — an ALE meta-analysis and MACM analysis. *Neuroimage* 87, 345–355. <https://doi.org/10.1016/j.neuroimage.2013.11.001>.
- Lahnakoski, J.M., Glerean, E., Jääskeläinen, I.P., Hyönä, J., Hari, R., Sams, M., Nummenmaa, L., 2014. Synchronous brain activity across individuals underlies shared psychological perspectives. *Neuroimage* 100 (100), 316–324. <https://doi.org/10.1016/j.neuroimage.2014.06.022>.
- Lamm, C., Meltzoff, A.N., Decety, J., 2010. How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 22 (2), 362–376. <https://doi.org/10.1162/jocn.2009.21186>.
- 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, 2492–2502. <https://doi.org/10.1016/j.neuroimage.2010.10.014>.
- Lovell, D., 1979. *Producer; Zeffirelli, F., Director. The Champ [Film]*, MGM/Pathé Home Video. Culver City, CA.
- Mar, R.A., 2011. The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol.* 62, 103–134. <https://doi.org/10.1146/annurev-psy-120709-145406>.
- Meffert, H., Gazzola, V., den Boer, J.A., Bartels, A.A.J., Keysers, C., 2013. Reduced spontaneous but relatively normal deliberate vicarious representations in psychopathy. *Brain* 136, 2550–2562. <https://doi.org/10.1093/brain/awt190>.
- Monfardini, E., Gazzola, V., Boussaoud, D., Brovelli, A., Keysers, C., Wicker, B., 2013. Vicarious neural processing of outcomes during observational learning. *PLoS One* 8 (9). <https://doi.org/10.1371/journal.pone.0073879> e73879.
- Morelli, S.A., Lieberman, M.D., Zaki, J., 2015. The emerging study of positive empathy. *Soc. Personal. Psychol. Compass* 9, 57–68. <https://doi.org/10.1111/spc3.12157>.
- Nanni, M., Martínez-Soto, J., Gonzalez-Santos, L., Barrios, F.A., 2018. Neural correlates of the natural observation of an emotionally loaded video. *PLoS One* 13 (6). <https://doi.org/10.1371/journal.pone.0198731> e0198731.
- Nastase, S.A., Gazzola, V., Hasson, U., Keysers, C., 2019. Measuring shared responses across subjects using intersubject correlation. *Soc. Cogn. Affect. Neurosci.* <https://doi.org/10.1093/scan/nsz037>.
- Nummenmaa, L., Glerean, E., Viinikainen, M., Jaaskelainen, I.P., Hari, R., Sams, M., 2012. Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci.* 109, 9599–9604. <https://doi.org/10.1073/pnas.1206095109>.
- Ochsner, K.N., Silvers, J.A., Buhle, J.T., 2012. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann. N. Y. Acad. Sci.* 1251, E1–E24. <https://doi.org/10.1111/j.1749-6632.2012.06751.x>.
- Pickett, C.L., Gardner, W.L., Knowles, M., 2004. Getting a cue: the need to belong and enhanced sensitivity to social cues. *Personal. Soc. Psychol. Bull.* 30, 1095–1107. <https://doi.org/10.1177/0146167203262085>.
- Pobric, G., Hamilton, A.F.D.C., 2006. Action understanding requires the left inferior frontal cortex. *Curr. Biol.* 16 (5), 524–529. <https://doi.org/10.1016/j.cub.2006.01.033>.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10 (2), 59–63. <https://doi.org/10.1016/j.tics.2005.12.004>.
- Preston, S.D., De Waal, F.B., 2002. Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25 (1), 1–20. <https://doi.org/10.1017/S0140525X02000018>.
- Raz, G., Winetraub, Y., Jacob, Y., Kinreich, S., Maron-Katz, A., Shaham, G., Podlipsky, I., Gilam, G., Soreq, E., Hendler, T., 2012. Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. *Neuroimage* 60 (2), 1448–1461. <https://doi.org/10.1016/j.neuroimage.2011.12.084>.
- Raz, G., Hendler, T., 2014. Forging Cinematic Paths to the Self: Neurocinematically Informed Model of Empathy in Motion Pictures. <https://doi.org/10.3167/proj.2014.080206>.
- Raz, G., Touroutoglou, A., Wilson-Mendenhall, C., Gilam, G., Lin, T., Gonen, T., Jacob, Y., Atzil, S., Admon, R., Bleich-Cohen, M., Maron-Katz, A., 2016. Functional connectivity dynamics during film viewing reveal common networks for different emotional experiences. *Cognit. Affect. Behav. Neurosci.* 16 (4), 709–723. <https://doi.org/10.3758/s13415-016-0425-4>.
- Reniers, R.L., Völlm, B.A., Elliott, R., Corcoran, R., 2014. Empathy, ToM, and self-other differentiation: an fMRI study of internal states. *Soc. Neurosci.* 9 (1), 50–62. <https://doi.org/10.1080/17470919.2013.861360>.
- Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>.
- Rizzolatti, G., Sinigaglia, C., 2010. The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11 (4), 264. <https://doi.org/10.1038/nrn2805>.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: the role of the temporoparietal junction in “theory of mind”. *Neuroimage* 19, 1835–1842. [https://doi.org/10.1016/S1053-8119\(03\)00230-1](https://doi.org/10.1016/S1053-8119(03)00230-1).
- Schaefer, A., Nils, F., Philippot, P., Sanchez, X., 2010. Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. *Cognit. Emot.* 24, 1153–1172. <https://doi.org/10.1080/02699930903274322>.
- Schnell, K., Bluschke, S., Konradt, B., Walter, H., 2011. Functional relations of empathy and mentalizing: an fMRI study on the neural basis of cognitive empathy. *Neuroimage* 54 (2), 1743–1754. <https://doi.org/10.1016/j.neuroimage.2010.08.024>.
- Schumann, K., Zaki, J., Dweck, C.S., 2014. Addressing the empathy deficit: beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *J. Personal. Soc. Psychol.* 107, 475–493. <https://doi.org/10.1037/a0036738>.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J., 2014. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* 42, 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>.
- Shaw, L.L., Batson, C.D., Todd, R.M., 1994. Empathy avoidance: forestalling feeling for another in order to escape the motivational consequences. *J. Personal. Soc. Psychol.* 67, 879–887. <https://doi.org/10.1037/0022-3514.67.5.879>.
- Simony, E., Honey, C.J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., Hasson, U., 2016. Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* 7, 12141. <https://doi.org/10.1038/ncomms12141>.
- Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162. <https://doi.org/10.1126/science.1093535>.
- Singer, T., Seymour, B., O’Doherty, J.P., Stephan, K.E., Dolan, R.J., Frith, C.D., 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469. <https://doi.org/10.1038/nature04271>.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, 2018–2209. <https://doi.org/10.1016/j.neuroimage.2004.07.051>.
- Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the

- brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U. S. A* 106, 13040–13045. <https://doi.org/10.1073/pnas.0905267106>.
- Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., Rizzolatti, G., 2001. I know what you are doing: a neurophysiological study. *Neuron* 31 (1), 155–165. [https://doi.org/10.1016/S0896-6273\(01\)00337-3](https://doi.org/10.1016/S0896-6273(01)00337-3).
- Valchev, N., Tidoni, E., Hamilton, A.F.D.C., Gazzola, V., Avenanti, A., 2017. Primary somatosensory cortex necessary for the perception of weight from other people's action: a continuous theta-burst TMS experiment. *Neuroimage* 152, 195–206. <https://doi.org/10.1016/j.neuroimage.2017.02.075>.
- Valk, S.L., Bernhardt, B.C., Trautwein, F.-M., Böckler, A., Kanske, P., Guizard, N., Collins, D.L., Singer, T., 2017. Structural plasticity of the social brain: differential change after socio-affective and cognitive mental training. *Sci. Adv.* 3 <https://doi.org/10.1126/sciadv.1700489> e1700489.
- Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L.C., Castellanos, F.X., 2015. Inscapes: a movie paradigm to improve compliance in functional magnetic resonance imaging. *Neuroimage* 122, 222–232. <https://doi.org/10.1016/j.neuroimage.2015.07.069>.
- Viinikainen, M., Jääskeläinen, I.P., Alexandrov, Y., Balk, M.H., Autti, T., Sams, M., 2010. Nonlinear relationship between emotional valence and brain activity: evidence of separate negative and positive valence dimensions. *Hum. Brain Mapp.* 31 (7), 1030–1040. <https://doi.org/10.1002/hbm.20915>.
- Weisz, E., Zaki, J., 2018. Motivated empathy: a social neuroscience perspective. *Curr. Opin. Psychol.* 24, 67–71. <https://doi.org/10.1016/j.copsyc.2018.05.005>.
- Westermann, R., Spies, K., Stahl, G., Hesse, F.W., 1996. Relative effectiveness and validity of mood induction procedures: a meta-analysis. *Eur. J. Soc. Psychol.* 26 (4), 557–580. [https://doi.org/10.1002/\(SICI\)1099-0992\(199607\)26:4<557::AID-EJSP769>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4).
- 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, 655–664. [https://doi.org/10.1016/S0896-6273\(03\)00679-2](https://doi.org/10.1016/S0896-6273(03)00679-2).
- Zaki, J., 2014. Empathy: a motivated account. *Psychol. Bull.* 140, 1608–1647. <https://doi.org/10.1037/a0037679>.