UMinho|2019

Sofia Raquel Pereira de Sousa Esménio

Understanding the neural basis of social cognition in a dyadic context with a close other using an effective connectivity approach

---

Universidade do Minho
Escola de Psicologia

Sofia Raquel Pereira de Sousa Esménio

Understanding the neural basis of social cognition in a dyadic context with a close other using an effective connectivity approach

junho de 2019

Universidade do Minho
Escola de Psicologia

Sofia Raquel Pereira de Sousa Esménio

Understanding the neural basis of social cognition in a dyadic context with a close other using an effective connectivity approach

Tese de Doutoramento em Psicologia Básica

Trabalho efetuado sob a orientação da
Doutora Joana Fernandes Pereira Coutinho
e do
Doutor José Miguel Montenegro Soares

junho de 2019

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Acknowledgements

I would like to thank my advisors, Dr. Joana Coutinho and Dr. José Miguel Soares for the astonishing time and attention they devoted to this project. Thank you Dr. Joana Coutinho for encouraging my autonomy, my critical thinking and inspire me to go further, reminding me that: “Only one who devotes himself to a cause with his whole strength and soul can be a true master. For this reason mastery demands all of a person.” Einstein. I would also like to thank Dr José Miguel Soares for always believing and supporting my ideas with enthusiasm.

A special thank you to Prof. Karl Friston and Dr. Adeel Razi, for welcoming me in your exceptional group. I am deeply grateful for the priceless knowledge shared and amazing availability. My stay in your lab was definitely a milestone in my career as a researcher. Thank you Peter Zeidman for being so helpful and accepting. For the infinite patience during our numerous talks, and for making me feel more like at home. Without you my stay would not have been as productive and enjoyable.

Thank you to everyone at the Psychological Neuroscience Laboratory for all the companionship, help, comments and suggestions. To my colleagues and friends, Margarida, Eugénia, Andreia, Álvaro, Nocas, Guida, Tatiana, Ricardo, Paulo, Carol, Fabiana, Diana, thank you so much for all the support, laughs and inspiring conversations. You have helped me make it till the end. Special thank you to Renata for always being there with wisdom, endless patience, and an amazing sense of humour. Meeting you made me a better researcher, but more importantly a better person.

To my dear friends João, Margarida, Nicole, Liliana, Madalena, Ângela, Filipa, Ricardo, José, I could never make it without your friendship, company, support and delightful humour. Thank you for believing in me even when I couldn’t anymore.

Finally, I would like to thank my close family for all the financial, emotional and physical support throughout this process. Special thank you to my parents, my godfather and wife Ana Viegas, João Encarnação and Dr. Mário Miguel Rosa without whom I would certainly not have made it.

Finally, thank you FCT for the funding that made this PhD possible (PD/BD/105963/2014).

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Understanding the neural basis of social cognition in a dyadic context with a close other using an effective connectivity approach

Humans interact with conspecifics on a daily basis in a multitude of social environments. In order to adjust one's own behaviour to the demands of a given social situation, individuals must be able to understand other's emotional and cognitive states. The main goal of the present dissertation was to explore the neural basis of social cognition in the context of a dyadic interaction with a significant other. Forty-two participants involved in a committed monogamous relationship underwent an fMRI session composed by a psychological assessment, a resting state acquisition and an ecological social cognitive task based on a couple´s interaction task previously performed in the laboratory. This task was combined with an effective connectivity (EC) approach to investigate the neural mechanisms involved in the capacity to infer our own and other's internal states.

In study 1 we looked at the functional integration within a well-known resting state network involved in social processing, the Default Mode Network (DMN), and explored how it relates to self-perceived social cognition. DMN's functional integration was assessed both in terms of functional connectivity and effective connectivity. Results show that although in Independent Component Analysis (ICA) higher scores in self-perceived empathy were associated with a significant increase in the medial prefrontal cortex (mPFC)'s expression, Dynamic Causal Modelling (DCM) clarified that this increase results mainly from a shift in coupling between posterior cingulate cortex (PCC) and right inferior parietal lobe (rIPL).

After measuring the brain's activity at rest, in study 2 we looked at the neural basis of self and other processing using a paradigm based on a real couples' interaction where participants were asked to elaborate on their partner´s experience (other-condition) or on their own experience (self-condition) while watching emotional video vignettes of their partner. The results revealed an almost complete brain overlap between the neural systems that process self and other's internal states. Despite this overlap the self-condition was also associated with brain regions involved in interoceptive and affect processing such as the posterior insula, whereas the other-condition recruited to further extent brain regions involved in mental inference and perspective taking like the temporarietal junction (TPJ).

Finally, after identifying the brain activation maps for self and other processing, in study 3 we used DCM to assess the functional integration among cognitive brain areas involved in self and other processing, in particular the bilateral TPJ, the PCC/precuneus and the bilateral middle temporal gyrus (MTG). DCM results showed that within this network first the sensory information is processed by the bilateral MTG, region well known for language and narrative comprehension. Only then the information follows simultaneously to the PCC, which plays a coordinative role within this network, and the rTPJ that will be more or less receptive depending if one is making inferences about others or the self, respectively.

Keywords: DCM; Romantic relationships; Self and other processing; Social cognition

Compreendendo as bases neurais da cognição social num contexto diádico com um outro significativo utilizando uma abordagem de connectividade efetiva

Os seres humanos interagem diariamente com outros indivíduos numa multiplicidade de contextos sociais. De modo a ajustar o seu comportamento às exigências de uma determinada situação social, os indivíduos devem ser capazes de compreender os estados emocionais e cognitivos dos outros. O objetivo principal desta dissertação foi o de explorar as bases neurais da cognição social no contexto de uma interação diádica com um outro significativo. Quarenta e dois participantes numa relação de casal estável foram submetidos a uma sessão de fMRI composta por uma avaliação psicológica, uma aquisição em estado de repouso e uma tarefa ecológica de cognição social baseada numa interação de casal realizada em laboratório. De modo a analisar os mecanismos neurais da capacidade de inferir os nossos estados internos e os de outros, uma abordagem de conectividade efetiva foi utilizada.

No estudo 1 analisamos a integração funcional na *Default Mode Network* (DMN), uma das mais bem estudadas redes cerebrais de descanso a qual está envolvida em diversos processos sociocognitivos, e explorámos a sua relação com a cognição social auto-percebida. Esta integração funcional foi avaliada em termos de conectividade funcional e efetiva. Os resultados mostraram que apesar de a análise de componentes independentes (ICA) mostrar que pontuações mais elevadas na escala de empatia diádica estavam associadas apenas a uma maior conetividade funcional do córtex medial prefrontal (mPFC), a modelagem causal dinâmica (DCM) permitiu clarificar que esse aumento resulta de uma mudança no acoplamento entre o cingulado posterior (PCC) e o lobo parietal inferior direito (rIPL).

Após medir a atividade do cérebro em repouso, no estudo 2, investigamos os circuitos cerebrais envolvidos no processamento do *self* e do outro utilizando uma tarefa social de fMRI que teve por base uma interação de um casal real. Nesta tarefa experimental os participantes elaboravam sobre a experiência do seu parceiro (condição *other*) ou sobre a sua própria experiência (condição *self)* enquanto assistiam um vídeo do seu parceiro a expressar conteúdos emocionais. Os resultados revelaram uma quase completa sobreposição anatómica e funcional entre os sistemas neurais subjacentes ao processamento do *self* e do *other*. Apesar disso, a condição *self* recrutou também áreas envolvidas no processamento interoceptivo e afetivo, como a ínsula posterior, enquanto a condição *other* recrutou ainda regiões envolvidas na inferência mental e tomada de perspetiva como a junção temporarietal (TPJ).

Finalmente, após identificar os mapas de ativação cerebral subjacentes ao processamento do eu e do outro, no estudo 3, usamos o DCM para avaliar a integração funcional entre áreas do cérebro envolvidas nestes dois processos, nomeadamente em áreas cognitivas tais como o TPJ bilateral, o PCC / precuneus e a circunvolução temporal média bilateral (MTG). Os resultados mostraram que nesta rede, a informação sensorial é primeiro processada pelo MTG, região de compreensão narrativa e da linguagem. Posteriormente, a informação segue conjuntamente para o PCC, o qual desempenha um papel coordenador dentro desta rede, e para a rTPJ, que estará mais ou menos receptivo dependendo se o individuo está a fazer inferências sobre os outros ou sobre o ele próprio, respetivamente.

Palavras-Chave: Cognição social; DCM; Processamento do *Self* e outro; Relações Romanticas

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **III**

ABSTRACT ..... **V**

RESUMO ..... **X**

ABBREVIATIONS, ACRONYMS AND SYMBOLS ..... **XI**

FIGURES ..... **XV**

TABLES ..... **XVIII**

CHAPTER I **1**

THEORETICAL BACKGROUND **1**

NEUROIMAGING APPROACHES TO THE STUDY OF PSYCHOLOGICAL PROCESSES 2

DYNAMIC CAUSAL MODEL 4

SOCIAL COGNITION AND ITS DIMENSIONS 8

THE NEED TO STUDY THE NEURAL BASIS OF SOCIAL COGNITION USING ECOLOGICAL RESEARCH PRADIGMS 11

THE PRESENT DISSERTATION 14

CHAPTER II **17**

"USING RESTING-STATE DMN EFFECTIVE CONNECTIVITY TO CHARACTERIZE THE NEUROFUNCTIONAL ARCHITECTURE OF EMPATHY" 17

ABSTRACT 18

INTRODUCTION 19

METHODS 22

Participants *22*

Empathy Measures *23*

MRI acquisitions *24*

Functional connectivity pre-processing and analysis *24*

Effective connectivity pre-processing and analysis *26*

RESULTS 27

Functional connectivity results *27*

Effective connectivity results *28*

DISCUSSION 29

Author Contributions *29*

Acknowledgements *29*

CHAPTER III 33

"BRAIN CIRCUITS INVOLVED IN UNDERSTANDING OUR OWN AND OTHER´S INTERNAL STATES IN THE CONTEXT OF ROMANTIC RELATIONSHIPS" 33

ABSTRACT 34

INTRODUCTION 35

METHOD 37

Participants *37*

Experimental Task *38*

Neuroimage acquisition *39*

Data preprocessing and analysis *40*

Statistical analysis *40*

RESULTS 41

Self Condition *41*

Other Condition *43*

Comparing Self/Other Condition *44*

DISCUSSION 46

Author Contributions *49*

Acknowledgements *49*

CHAPTER IV 50

"CHANGES IN THE EFFECTIVE CONNECTIVITY OF THE SOCIAL BRAIN WHEN MAKING INFERENCES ABOUT SELF AND CLOSE OTHERS" 50

ABSTRACT 51

INTRODUCTION 52

METHOD 55

Participants *55*

Experimental Task *55*

Image acquisition *56*

Data preprocessing *56*

Effective connectivity analysis *57*

RESULTS 59

DISCUSSION 61

Author Contributions *61*

Acknowledgements *64*

CHAPTER V 65

DISCUSSION AND CONCLUSIONS 65

REFERENCES 72

APPENDICES 95

APPENDIX A 96

Supplementary Material – Introduction 96

APPENDIX B 98

Supplementary Material – Chapter IV 98

ABBREVIATIONS, ACRONYMS AND SYMBOLS

AAL - Anatomical Automatic Labeling atlas

ACC - Anterior cingulate cortex

AI – Anterior Insula

AMCC - Anterior midcingulate cortex

ANOVA – Analysis of Variance

BMR – Bayesian Model Reduction

BOLD – Blood-oxygen-level dependent

CSD – Cross spectra density

CSF - Cerebrospinal fluid

CVA - Canonical Variance Analysis

DAS - Dyadic adjustment scale

DCM – Dynamic causal modelling

DEC - Dyadic Empathic Concern

DeoxyHb - Deoxyhaemoglobin

DMN – Default Mode Network

DPT - Dyadic Perspective-Taking

EC – Effective connectivity

EP- Parameters posterior estimates

EPI - Echo-planar imaging

FC – Functional connectivity

fMRI – Functional magnetic ressonance imaging

FoV - Field of view

FSL - FMRIB Software Library

FWE – Family wise error

GLM - General linear model

HRF - Hemodynamic Response function

ICA – Independent Component Analysis

IRI - Interpersonal Reactivity Index

IRIC - Interpersonal Reactivity Index for Couples

IFG – Inferior frontal gyrus

IPL - Inferior parietal lobe

LIPL – Left inferior parietal lobe

LMTG – Left middle Temporal gyrus

LTPJ – Left tempoparietal junction

MNI - Montreal Neurological Institute

mPFC - Medial prefrontal cortex

MNI - Montreal neurological institute

MRI - Magnetic ressonance imaging

MTG – Middle temporal gyrus

MT/V5 - Middle temporal visual area

OxyHb - Oxyhaemoglobin

PCC – Posterior cingulate cortex

Prec - Precuneus

PEB – Paramerical Empirical Bayes

pSTS - Posterior superior sulcus

PT – Perspective Taking

RFX - Random effects

RIPL – Right inferior parietal lobe

RMTG – Right middle Temporal gyrus

ROI – Region of interest

RTPJ – Right tempoparietal junction

SD – Standard Deviation

SMA - Supplementary motor area

SMG - Supramarginal gyrus

spDCM – Spectral Dynamic causal modelling

SPM - Statistical Parametric Mapping software

STG – Superior temporal gyrus

STP – Superior temporal poles

STS - Superior temporal sulcus

SVC – Small volume correction

TOM – Theory of Mind

TE – Echo time

TPJ - Temporoparietal junction

TR – Repetition time

WM – White matter

FIGURES

Chapter I

Figure I – 1 Possible models of functional integration between regions, based on similar BOLD signal correlations. Adapted from Kahan & Foltynie (2013).

Figure I- 2 Scheme showing the DCM underlying equations. The data contain observation noise $\varepsilon(t)$ that is added to regional observations. Adapted from (Friston, 2009a).

Figure I - 3 Haemodynamics model (a) Hemodynamic Response function (HRF) scheme (b) Single region haemodynamic model: an increase in neuronal activity $z(t)$ induces a signal (s) that stimulates the blood flow (f), changing the volume (v) and deoxyhemoglobin (q); which then enter an output to give the observed BOLD signal (y). Adapted from (Friston, 2009a).

Figure I - 4 Simplified version of study 3 fMRI experiment combining only 3 regions of interest (ROIs). In DCM, parameters $a_{k,j}$ represents the EC from node *j* to *k*; parameters $c_{k,u}$ denote the effect of the driving input *u* on target node *k*; and parameters $b_{k,u}$ denote the modulatory effect of the input *u* on target connection *k*. Adapted from Kahan & Foltynie (2013).

Figure I- 5 Scheme of one trial of the implemented fMRI task.

Figure I - 6 Schematic representation of the interaction task performed in laboratory. The selected sequence had in consideration ethical and methodological options of other marital interaction studies (e.g., Perrone-McGovern et al., 2014).

Chapter II

Figure II - 1 a) Consistent DMN expression identified in resting state over subjects. b) Correlation between IRIC total score and DMN functional connectivity.

Figure II – 2 a) Canonical Vector containing weight (contribution) assigned by CVA to each DCM effective connectivity parameter. Blue connections correspond to extrinsic connections and green to intrinsic connections. b) Canonical Vector containing weight (contribution) assigned by CVA to each empathy scale score: perspective taking (PT) and empathic concern (EC). c) Showing the index number of each effective connection in (a) d) weight of each connection, where bold lines represent an increase and dashed lines a decrease with empathy.

Chapter III

Figure III - 1 Scheme of an emotional and a neutral trial in the self-condition.

Figure III - 2 - Activations identified in the self-condition when hearing to a) positive (versus neutral) contents or b) negative (versus neutral) contents expressed by their partner.

Figure III – 3 Activations identified in the other-condition when hearing to a) positive (versus neutral) contents or b) negative (versus neutral) contents expressed by their partner.

Figure III - 4 Activations identified when hearing to contents expressed by their partner for a) self-condition higher than the other condition or b) other-condition higher than the self-condition

Chapter IV

Figure IV - 1 - Scheme of a self and other emotional trial.

Figure IV - 2 – DCM initial model. a) Connectivity architecture. b) Driving inputs. c) Modulatory effects. D) Final model.

Figure IV - 3 - "Baseline" Connectivity results. a) Parameters posterior estimates (EP). b) Connectivity matrix. c) The structure and parameters of the winning model. The black lines/values illustrate the (natural) connectivity between brain regions, i.e. irrespective of stimulus and task. The numbers are the strength of connectivity (Hz).

Figure IV - 4 – Driving inputs and modulatory effect results. a) Driving inputs. b) Modulatory effects or condition-specifics. (left) Parameters posterior estimates (EP). (center) Connectivity matrix. (right) The structure and parameters of the winning model. The black lines/values illustrate the (natural) connectivity between brain regions. The arrows in blue represent the driving inputs (up) and modulatory effects (down), respectively. The numbers are the strength of connectivity (Hz).

Appendix A

Supplementary Figure I - 1 DCM fMRI experiment in study 3. Parameters $a_{k,j}$ represents the EC from node j to k; $c_{k,u}$ denote the effect of the driving input u on target node k; and $b_{k,u}$ denote the modulatory effect of the input u on target connection k. Adapted from Kahan & Foltynie (2013).

Supplementary Figure I – 2 Scheme of the DCM resting state fMRI experiment conducted in study 1. Coupling parameters $a_{k,j}$ represents the EC from node j to k. Adapted from Kahan & Foltynie (2013).

Appendix B

Supplementary Figure IV - 1. Results of specifying a modulatory effect in every intrinsic (self-) connection. (a) Parameters posterior estimates (EPs). (b) Connectivity matrix. (c) Structure and parameters. The black lines/values illustrate the (natural) connectivity between brain regions. The arrows in blue represent the modulatory effects. The numbers are the strength of connectivity (Hz).

TABLES

Chapter III

Table III - 1. Regional brain activity in the self condition for positive (> neutral) and negative (> neutral) stimuli.

Table III - 2. Regional brain activity in the other condition for positive (> neutral) and negative (> neutral) stimuli.

Table III - 3. Differences in brain activation between self and other conditions; Voxel Threshold: p <= .005; Cluster Size Threshold: k >15.

Table III - 4. Differences in brain activation for positive (> negative) contents expressed by their partner; Voxel Threshold: p <= .001; FWE corrected; Cluster Size Threshold: k >15.

CHAPTER I

THEORETICAL BACKGROUND

Neuroimaging approaches to the study of psychological processes

In the last decades a broad range of psychological processes has been studied by neuroscience researchers, ranging from basic ones such as perception or attention, to more complex and inherently human processes such as our capacity to reason and understand others' emotions. This exponential increase in cognitive, social and affective neuroscience was also the result of the development of different methodologies including imaging techniques. Soon after the first successful experiments in 1992 describing the injection of a susceptibility contrast agent to map blood volume in humans (Bandettini, Wong, Hinks, Tikofsky, & Hyde, 1992; Ogawa et al., 1992), neuroimaging studies particularly structural and functional magnetic resonance imaging (fMRI) progressively became one of the most used techniques for the study of the neural basis of a wide range of psychological processes.

Despite its relatively recent history, functional neuroimaging research has changed some of its traditional assumptions. In fact in the first decades of neuroimaging it was thought that neural processing consisted exclusively on neuronal activity in discrete points throughout the brain, more specifically in spatially distributed regions that were specialised in processing a given stimuli or behaviour (Friston, 1994). This principle has been referred to as "functional segregation" or localizationism. The idea that there was a linear relationship between distinct brain regions and specific functions dominated the scientific thinking about how the brain functioned in the 19th century. Following on Gall's phrenology findings (Gallup, Frederick, & Nathan Pipitone, 2008) and clinical experiments by Broca and Wernicke (Broca, 1865; Wernicke, 1874) showing that patients with focal brain lesions had specific language impairments, brain imaging has firmly established functional segregation as a principle of brain organization in humans (Friston, 2004, 2011).

Functional segregation's main assumption is that a cognitive function can be localized in a specific cortical area. This assumption has nevertheless been called into question because not only it is difficult to attribute a specific function to a single brain region, but also cerebral activity is highly dependent on the communication between separate brain regions (Friston, 2002; Hohwy, 2007). Instead, these brain regions (or modules) are in constant communication with each other sending information back and forward. This concept is known as "functional integration" or connectionism and suggests that the brain regions supporting a single function can be anatomically separated (Friston, 2004, 2011).

Thus in the beginning of the 21st century the attention of brain function research started shifting from functional segregation to functional integration (Friston, 2011). The functional integration among brain regions can be described either in terms of functional connectivity (FC) or effective connectivity (EC) (Friston, 1994; Horwitz, Tagamets, & McIntosh, 1999). Both measures depart from the fMRI signal of different brain areas, however FC is inferred on the basis of statistical dependencies among measurements of neuronal activity, i.e. correlations (Rogers, Morgan, Newton, & Gore, 2007); whereas effective connectivity measures the influence that one neural system exerts over another (Friston, 1994).

Important contributions to our understanding of the neural basis of different psychological processes have been made by neuroimaging studies assessing functional integration in terms of functional connectivity (e.g. Chen, Ye, Jin, Zhu, & Wang, 2019; Kim et al, 2017; Oliveira Silva et al., 2018; Tomasi & Volkow, 2012; Whitfield-Gabrieli & Ford, 2012). Nevertheless, despite its important contribution, FC only measures which voxels in the brain display similar blood-oxygen-level dependent (BOLD) signal fluctuations over time, it does not consider how the information propagates through those brain regions (Friston, 2011b). In other words, FC does not allow the characterization of the communication patterns among the brain regions engaged in a specific task. These regions can either communicated in several different ways, or not communicate at all (Figure I-1).

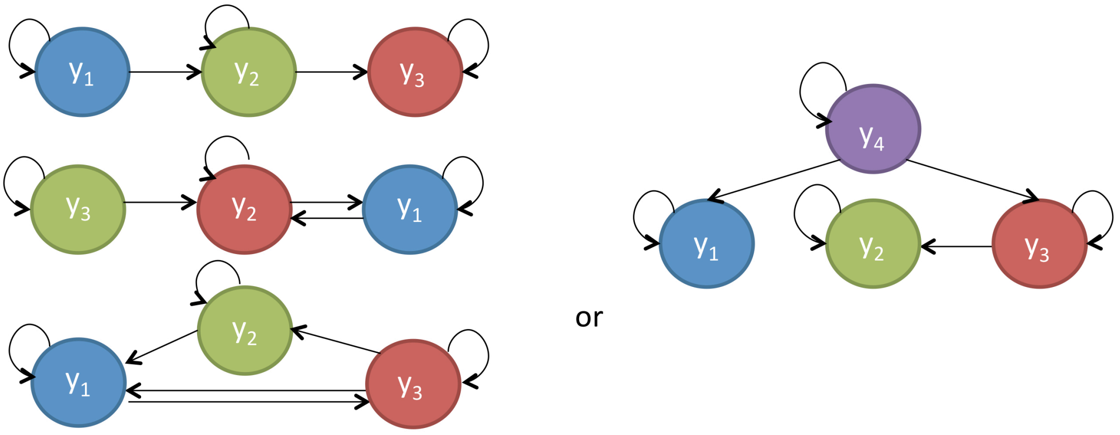


Figure I -1 Possible models of functional integration between regions, based on similar BOLD signal correlations. Adapted from Kahan & Foltynie (2013).

Since correlations can arise in numerous ways, functional integration within a network is better understood using EC methods that specifically measure how the activity in one brain region influences the activity in another brain region (Breakspear, 2004; Stephan & Friston, 2010). In essence, EC methods are dynamic and dependent on a model of interactions or coupling that explains how the observed signal was generated, i.e. a “generative model” (Aertsen & Preibl, 1991; Friston, 1994). As such, instead of simply analysing the signal, EC methods consider how the signal was generated. This is the foundation of Dynamic Causal Modelling (DCM), which is the EC method used in study 1 and study 3 of the present dissertation to unravel specific patterns of internode communication (Friston, Harrison, & Penny, 2003).

Dynamic Causal Model

DCM is a quite recent hypothesis-driven Bayesian approach that has become the gold standard method for modelling causal interactions in neuroimaging data (e.g. Bastos-Leite et al., 2015; Hyett et al., 2015; Tang, Razi, Friston, & Tang, 2016). In essence, DCM infers EC between brain regions using dynamic neuronal models of how cortical regions interact, combined with a detailed biophysical haemodynamic model that transforms neuronal activity into the measured response (Buxton & Frank, 1997; Daunizeau, David, & Stephan, 2011; Friston et al., 2003; Stephan & Friston, 2010).

These neuronal causal models, which explain the “hidden” communication between the regions of interest, are defined within the general mathematical framework provided by dynamic systems theory (Breakspear, 2004; Jirsa, 2004). In particular, in fMRI, the neuronal models are based on fairly simple equations describing the evolution of neuronal states (Figure I- 2).

$$y(t) = g\big(z(t)\big) + \varepsilon(t)$$

$$\begin{bmatrix} \dot{z}_1 \\ \cdots \\ z_N \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{bmatrix} + u(t) \begin{bmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots & b_{NN} \end{bmatrix} \begin{bmatrix} z_1 \\ \cdots \\ z_N \end{bmatrix} + u(t) \begin{bmatrix} c_1 \\ \cdots \\ c_N \end{bmatrix}$$

Figure I- 2 Scheme showing the neuronal equations. Adapted from Friston (2009a).

where $z(t)$ describe the changes in neuronal activity; $u(t)$ the experimental manipulation or 'input'; $a_{k,j}$ represents the EC from node *j* to *k*; c_S is the effect of the driving input u_S ; and $b_{k,S}$ refers to a modulatory effect of the input u_S.

The neuronal activity $z(t)$ then enters a haemodynamic model that converts it into predicted haemodynamic signals $y(t)$. More specifically, the hemodynamic response function (HRF) transforms the neuronal states into haemodynamic response through a convolution function (Figure I – 3).

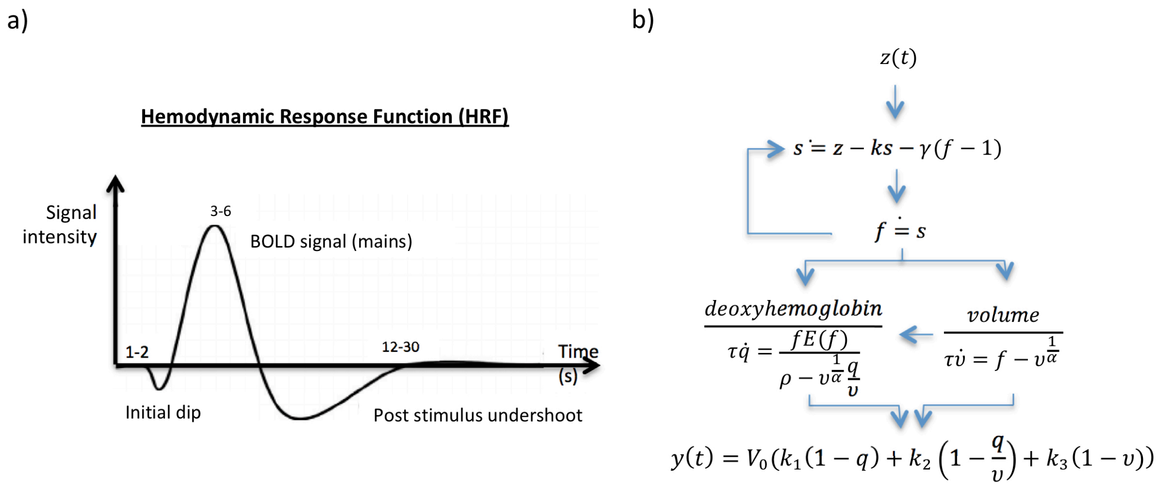


Figure I- 3 Haemodynamics model (a) Hemodynamic Response function (HRF) scheme (b) Single region haemodynamic model: an increase in neuronal activity z(t) induces a signal *(s)* that stimulates the blood flow *(f)*, changing the volume (*v)* and deoxyhemoglobin (*q*); which then enter an output to give the observed BOLD signal (*y)*. Adapted from Friston (2009a).

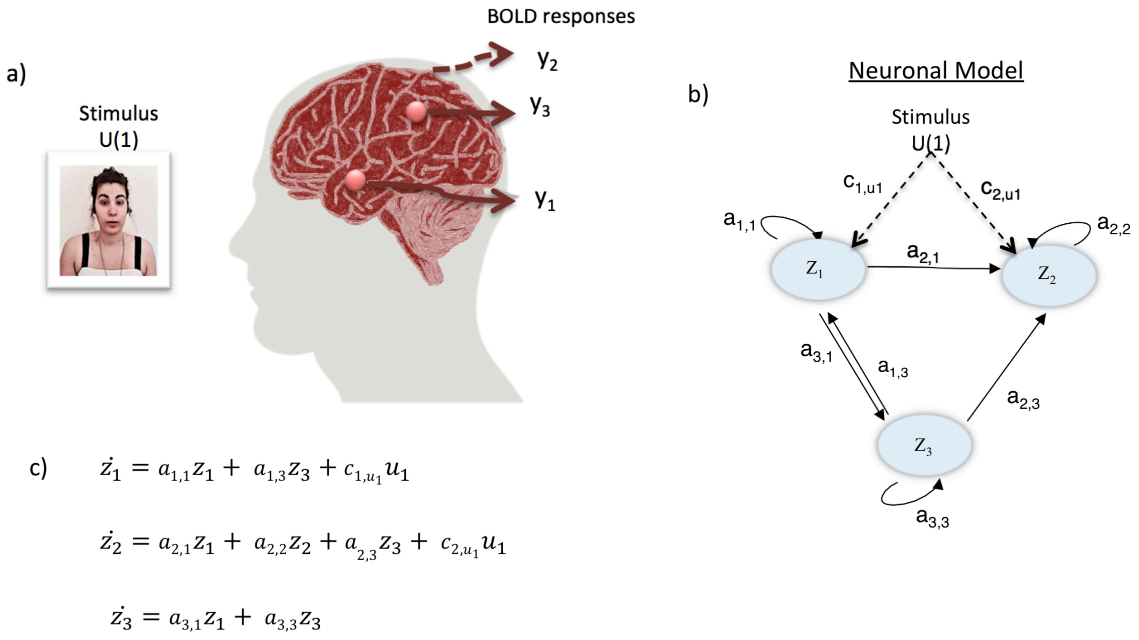
In order to comprehend the haemodynamic model in fMRI, it is important to understand how the observed BOLD is generated. Essentially, changes in BOLD signal are due to differences in the magnetic properties of deoxyhaemoglobin (deoxyHb) and oxyhaemoglobin (oxyHb). More specifically, the OxyHb is diamagnetic, meaning that it is innocuous to the magnetic field, whereas the deoxyHb is paramagnetic leading to a reduction in the local magnetic field induced by local dephasing of protons (Pauling & Coryl 1936). These magnetic differences combined with changes in blood relative composition induced by the haemodynamic response to neural activity, underlie the observed BOLD signal fluctuations.

In brief, when a brain region is recruited during an fMRI task there is an initial reduction in oxyHb and increase in deoxyHb due to local capillaries extraction that causes an initial dip in BOLD signal. This dip is then followed by an increase in cerebral blood flow that delivers oxyHb overshadowing oxygen consumption by a 2:1 rate (Buxton & Frank, 1997). This large rebound in local oxygenation leads to an increase in the local BOLD signal with a lag of 2-6 seconds; signal that can be imaged using a T2* weighted sequence. As such, in fMRI DCM the BOLD signal represents the observable variable (y) of the neural activity (z), which is the "hidden state variable". In DCM these neural states (z) that generate the BOLD data (y) are sensitive to afferents from other regions and exogenous inputs (Friston, Kahan, Biswal, & Razi, 2014; Razi, Kahan, Rees, & Friston, 2015).

As represented in Figure I – 2 and Figure I – 3, both the dynamic neuronal model and the biophysical haemodynamic model are specified in continuous time in terms of non-linear differential equations. These equations contain parameters that encode the strength of connections (and how they change with manipulation) and other biophysical rate constants (Daunizeau et al., 2011; Stephan & Friston, 2010; Zeidman et al., 2019). These are the parameters that DCM tries to optimize.

The optimisation of these parameters provides the set of parameters that most likely explains the model's evidence as well as the probability of observing that data under a particular model, i.e. model evidence. Specifically, the model evidence enables the comparison of different models and consequent selection of the one that better explains the data. Finally, if each model is associated with a neuronal architecture hypothesis, the model evidence also provides a quantitative measure for comparing different theoretical hypotheses.

To illustrate the application of DCM to an fMRI task, a simplified version of the study 3 experimental setting is displayed in Figure I – 4. The full experimental setting is displayed in Supplementary Figure I – 1 in Appendix A. Furthermore, for a complete guide to DCM first level analysis for fMRI see Zeidman et al. (2019).



$$\dot{z}_1 = a_{1,1}z_1 + a_{1,3}z_3 + c_{1,u_1}u_1$$

$$\dot{z}_2 = a_{2,1}z_1 + a_{2,2}z_2 + a_{2,3}z_3 + c_{2,u_1}u_1$$

$$\dot{z}_3 = a_{3,1}z_1 + a_{3,3}z_3$$

Figure I – 4 Version of study 3 fMRI experiment combining only 3 regions of interest (ROIs). In DCM, parameters $a_{k,j}$ represents the EC from node j to k; parameters $c_{k,u}$ denote the effect of the driving input u on target node k; and parameters $b_{k,u}$ denote the modulatory effect of the input u on target connection k. Adapted from Kahan & Foltynie (2013).

Finally, recent developments in DCM - in particular the creation of spectral DCM (spDCM) - now allow the modelling of the intrinsic dynamics in a resting state network (Friston et al., 2014; Razi et al., 2015). Essentially, spectral DCM estimates the EC among brain regions in the frequency domain, avoiding the estimation of random fluctuations in neural states. In other words, spDCM replaces the original BOLD timeseries with their second-order statistics - their cross spectra – which means that instead of estimating time varying hidden states, spDCM estimates their covariance that does not change with time. Essentially, spDCM models the endogenous activity that underlies the FC observed (Friston et al., 2014; Razi et al., 2015).

In resting state DCM a stochastic term to model endogenous neuronal fluctuations (υ) is added to the standard DCM neural equations (Figure I- 4). Considering that in resting state there are no task-related modulatory effects or driving inputs, the neural equation becomes: $\dot{z} = Az + \upsilon$, where A is the Jacobian describing the EC of the system in the absence of the fluctuations υ. The application of DCM to resting state is exemplified in Supplementary Figure I – 2 in Appendix A, which illustrates the resting state fMRI experiment of study 1.

Finally, the comparison of the simplified and full version of study 3 (Figure I – 4 and Supplementary Figure I – 1), clearly illustrates how the amount of connections between nodes increases in proportion with the augment of the areas recruited by a specific process, which leads to a greater number of plausible explanations of how the information flows between these areas. This increase in the number of model structures that can explain similar BOLD correlations is the main reason why effective connectivity methods are more appropriate when studying the neural basis of complex psychological functions that engage the cooperation of numerous brain regions, such as human empathy and theory of mind (e.g. Bzdok et al., 2012; Kanske, Böckler, Trautwein, & Singer, 2015; Reniers, Völlm, Elliott, & Corcoran, 2014; Völlm et al., 2006; Winter, Spengler, Bermpohl, Singer, & Kanske, 2017). These high-level functions are central in social cognition, which constitutes the focus of the present dissertation.

Social cognition and its dimensions

Social cognition refers to the processing of stimuli relevant to understand social agents and their interactions, encompassing complex psychological processes crucial for social adjustment (Chen, Martínez, & Cheng, 2018). In other words, social cognition refers to the ability to understand and make sense of ourselves, others and interpersonal relationships, allowing us to navigate in an otherwise unpredictable social world. Interestingly, evolutionary psychology and neuroscience have shown the existence of social abilities across different species such as primates, corvids and mammals (e.g. Bugnyar, Reber, & Buckner, 2016; de Waal, 2011; Decety, 2010; Wharton, Basu, & Ashe, 2005, for a quick review see Chen et al., 2018); pointing to the central role of social behaviour for the survival of individuals in large social groups.

Most of the current theoretical accounts on social cognition see it as a multidimensional construct (e.g. Alcalá-López et al., 2017; Decety & Lamm, 2006; Reniers et al., 2014; Shamay-Tsoory, 2011; Singer, Critchley, & Preuschoff, 2009), although a considerable debate exists between authors in terms of the what are the main domains of social cognition, as well as whether they constitute distinct or interrelated processes (Happé, Cook, & Bird, 2016).

In fact, several domains of social cognition have been identified, ranging from more basic dimensions such as social attention or encoding of social stimuli, to higher-order social processes such as social decision-making or social inference (Fiske & Taylor, 2013). Some of the most studied dimensions of social processing include: 1) social attention - the degree of attention paid to social stimuli (e.g. Freeth, Foulsham, & Kingstone, 2013; Salley & Colombo, 2016); 2) emotional processing and empathy - the ability to share and experience the affective state of self or another (e.g. Carr et al., 2003; Singer et al., 2004, 2006), 3) biological motion detection and action perception (e.g. Baetens, Ma, Steen, & Van Overwalle, 2013; Gallese & Sinigaglia, 2011); 4) mental state attribution and theory of Mind (TOM) - the ability to identify and understand self and other's mental states and anticipate his/ her actions (e.g. Frith & Frith, 2006; Saxe, Carey, & Kanwisher, 2004); 5) self-other distinction – the ability to properly distinguish between self and other representations (Coutinho, Silva, & Decety, 2014; Reniers et al., 2014).

Over the last two decades, social neuroscience research has helped to unravel the neural basis of these different dimensions of social cognition (e.g. Bzdok et al., 2012; de Vignemont & Singer, 2006; Decety & Lamm, 2006; FeldmanHall, Dalgleish, Evans, & Mobbs, 2015; Frith, 1999; Schilbach et al., 2012; Overwalle & Baetens, 2009). Two main routes of social processing involved in understanding others' and our own internal states have been the focus of this research agenda: an affective route, that enables us to vicariously experience and share others' affective states (e.g. de Vignemont & Singer, 2006), and a cognitive route involved in the inference of other´s mental states such as thoughts, beliefs or emotions (e.g. Frith & Frith, 2005, 2006).

The first route or socio-affective dimension corresponds to embodied or bottom up dimensions of social processing and has been commonly referred in the literature by emotional contagion, empathy or emotional empathy (Decety & Lamm, 2006; Shamay-Tsoory, 2011; Singer & Lamm, 2009). The second route is a more representational or top down path of processing that involves higher cognitive functions and is often termed in the literature as Theory of Mind (TOM), mentalizing or cognitive empathy (Carrington & Bailey, 2009; Frith & Frith, 2005; Lamm, Batson, & Decety, 2007; Saxe & Kanwisher, 2003; Saxe & Powell, 2006). These routes recruit two independent brain networks which will be described below (Eres, Decety, Louis, & Molenberghs, 2015; Kanske et al., 2015; Lamm et al., 2011; Singer, 2006).

The affective route of social processing involves the vicarious experience of another person´s feelings or inner states. This process through which the observer shares the other person's affect, leads to isomorphic emotional states in the experiencer and the observer (de Vignemont & Singer, 2006). Therefore, this affective dimension recruits a brain network similar to the one recruited in the first hand emotional experience (e.g. Botvinick et al., 2005; Decety, Michalska, & Akitsuki, 2008; Novembre, Zanon, & Silani, 2015; Wicker et al., 2003). The key regions of this affective network include the anterior cingulate cortex (ACC), anterior midcingulate cortex (aMCC), supplementary motor area (SMA), the anterior insula (AI) and the inferior frontal gyrus (IFG) (Bzdok et al., 2012; Fan, Duncan, de Greck, & Northoff, 2011; Shamay-Tsoory, 2011; Singer & Lamm, 2009). In addition, depending on the specificities of the experimental task, other regions may be recruited such as the posterior cingulate (PCC), anterior thalamus, inferior parietal lobe (IPL -including the supramarginal gyrus (SMG) and the angular gyrus) and midbrain (e.g. Kanske et al., 2015; Lamm, Meltzoff, & Decety, 2010; Lang, Yu, Markl, Müller, & Kotchoubey, 2011; Meyer et al., 2013).

On the other hand, the cognitive route of social processing refers to the ability to adopt another person's cognitive perspective in order to recognize and understand others' mental states (Frith & Frith, 2005; Frith & Frith, 2003; Saxe, Carey, & Kanwisher, 2004; Saxe & Kanwisher, 2003). In particular, according to Frith and Frith (2003), mentalizing is the process by which we make inferences about mental states, either those of others or ours. This capacity to "read the minds" of others has been associated with enhanced activity in posterior/ventral tempoparietal junction (TPJ), precuneus, middle temporal gyrus (MTG), middle temporal visual area (MT/V5) and a network including superior temporal sulcus (STS), temporal poles and medial prefrontal cortex (mPFC), specifically the ventromedial and dorsalmedial portion (Alcalá-López et al., 2017; Kanske et al., 2015; Mitchell, 2009; Schurz et al., 2014).

Similarly to what happens in the affective route in which the networks of self and other experience overlap, neuroimaging research has shown that the cognitive route also recruits similar brain systems when making inferences about self and others (Frith, 1999; Lombardo et al., 2009; Mitchell, Macrae, & Banaji, 2006; Saxe, Moran, Scholz, & Gabrieli, 2006). Since the inner states of others are not readily available, it seems that observers use their own emotional and mental states as the basis for understanding those of others. Indeed, neuroimaging evidence supports this idea that in the process of understanding others, one relies on psychological and neural mechanisms similar to those used to understand ourselves (Böckler, Herrmann, Trautwein, Holmes, & Singer, 2017; Gallese, 2014; Lamm, Bukowski, & Silani, 2016; Lamm et al., 2011).

This level of functional overlap between self and other processing will be the focus of studies 2 and 3 of this dissertation. As we will discuss in the next section, these studies intended to overcome the lack of ecological validity of most studies found in the literature.

The need to study the neural basis of social cognition using ecological research pradigms

Despite the significant advances in our understanding of the neural circuits involved in social processing, very few studies look at the brain´s response in more ecological scenarios, relying instead in stimuli that differ qualitatively from the kind of information one must process in real-life social encounters (Zaki & Ochsner, 2009). Most studies in social neuroscience claim to investigate the brain basis of social processes, however they most often use experimental tasks where subjects are exposed to simplified stimuli or fictional targets (e.g. Brass, Bekkering, & Prinz, 2001; Saarela et al., 2006). According to some authors this lack of ecological validity when studying social processes can result in artificial and simplistic ideas about complex psychological constructs (Schilbach, 2014; Schilbach et al., 2013). Additionally, these authors point to the current need of investigating the neural basis of the active participation in an online social interaction (which captures the interactor's point of view), instead of measuring the neural activity during passive participation (which captures the observer's point of view). In line with these ideas, one of the main goals of this dissertation was to develop a more ecological fMRI paradigm (described later in this section) resembling a real life interaction with a close other in particular a romantic partner.

Romantic relationships were the dyadic context that we selected to develop our studies. Among the various human interactions, romantic relationships constitute a good model to analyse the neural mechanisms of self and other processing considering the emotional proximity with the target. According to some authors, being in an intimate relationship can be described in terms of including the other in the self (e.g. Aron, Aron, & Smollan, 1992; Aron, Aron, Tudor, & Nelson, 1991). Indeed, the proximity and emotional closeness between members of a couple is likely to be associated with the creation of overlapping cognitive structures between the self and the partner (Péloquin & Lafontaine, 2010; Waldinger et al., 2004). Supporting this idea, empirical evidence showed that imagining oneself, a loved-one and a stranger in pain resulted in different degrees of overlap in the pain matrix, which were related to the participants' subjective reports of closeness (Cheng et al., 2010).

Our choice of couples as a context of study was also justified by the clinical and social relevance of this research topic. Research has shown that while relationship satisfaction is associated with partners’ psychological health, marital conflict has a well-documented negative impact on the physical (Bookwala, 2005; Gottman & Levenson, 1988; Lehnart, Neyer, & Eccles, 2010; Robles & Kiecolt-Glaser, 2003) and mental health of each partner (Beach, Fincham, & Katz, 1998; Gabriel, Beach, & Bodenmann, 2010; Rehman, Gollan, & Mortimer, 2008). The study of the mechanisms involved in couple´s interactions has clear implications for the understanding of marital conflict and its social, familiar and labour consequences (Cummings & Davies, 2002; El-Sheikh et al., 2009).

Overall, these findings emphasise the importance of assessing social ability in a dyadic context. However, as mentioned before, studies addressing the brain’s response in more realistic interpersonal interactions are still scarce. The use of ecological tasks is particularly important for the study of social cognition and social interactions, because in naturally occurring social encounters there are several factors that influence the neural response to other´s emotions, such as emotional context (Han et al., 2009); perceived fairness (Singer et al., 2006); familiarity (Meyer et al., 2013); intimacy (Cheng, Chen, Lin, Chou, & Decety, 2010); similarity and need for care (Batson, Lishner, Cook, & Sawyer, 2005).

Thus, in the present work, we used the context of a meaningful relationship with a significant other (the romantic partner) to investigate the neural mechanisms involved in social cognition. More specifically, we used an fMRI paradigm based on an interaction task that resembles couples’ natural interactions (Mayne et al., 1997) to study the mechanisms involved in participants´ ability to understand their own and their partner´s inner states. In sum, in this task each participant watched a series of videos-vignettes of his/her romantic partner expressing emotional contents and was asked to focus on his/her spouse´s experience (other condition) or on his/her own experience (self condition).

This fMRI paradigm used a standard block design with a 2 x 3 factorial structure that included a self-other factor and three levels of valence (i.e. positive, negative and neutral). Each block contained 22 trials: eight positive videos, i.e. things that their partner liked about the relationship or admire about them; eight negative videos, i.e. things their partner disliked about them, or they would like to change, and six neutral videos of landscapes extracted from the Emotional Movie Database (Carvalho, Leite, Galdo-Álvarez, & Gonçalves, 2012). Each trial started with a fixation cross, followed by instructions in accordance with the present block. Example of a single trial is provided in Figure I - 5.

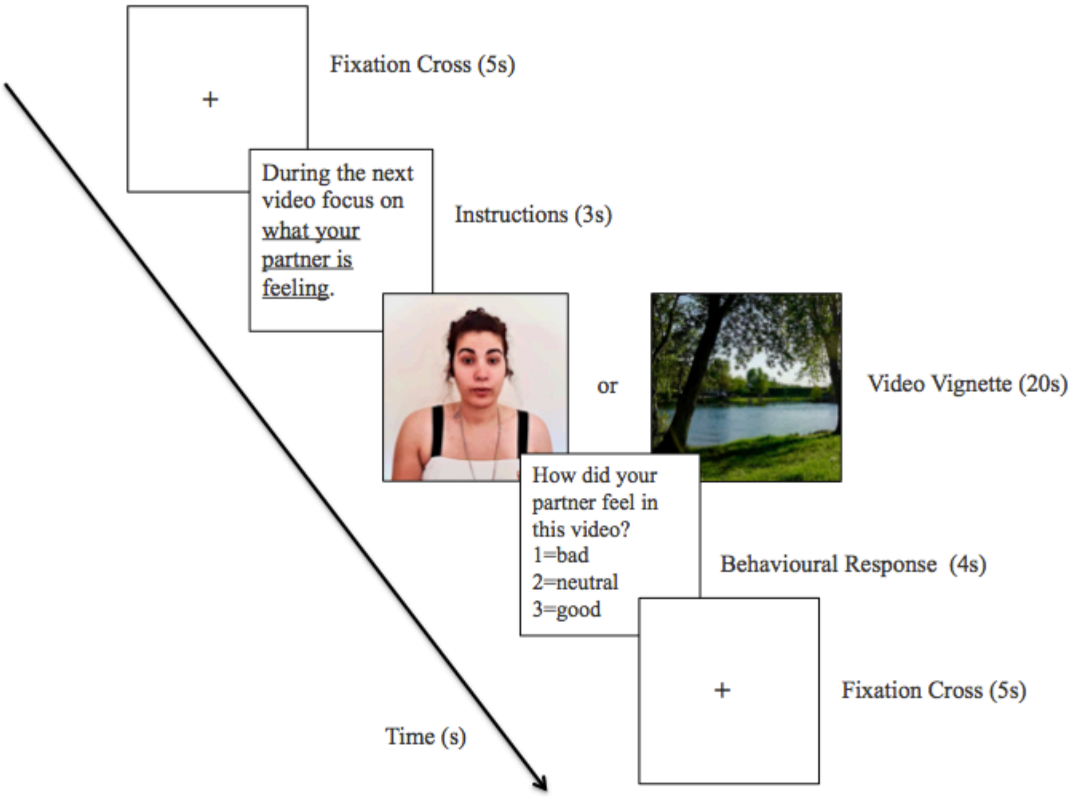


Figure I- 5 Scheme of one trial of the implemented fMRI task.

It is important to mention that the video-vignettes used as stimuli in this fMRI task were extracted from a real interaction that had been previously performed by the same participants in the lab. This previous interaction consisted of a structured discussion where both spouses talked about negative and positive aspects of their relationship. More specifically, each member of the couple was instructed that they had two minutes to talk to their partner about a given negative aspect at a time. They should talk with detail about each topic, by giving specific examples or by talking about the impact that it had on them. The other partner could not argue or answer back, but should only paraphrase what he/she had just heard. Partners switched their roles until they had covered all the important topics. The topics to be discussed could be chosen from a list of common issues in marital relationships, or be picked by the participant. The same procedures and instructions were applied to the following discussion of the positive aspects as shown in Figure I -6.

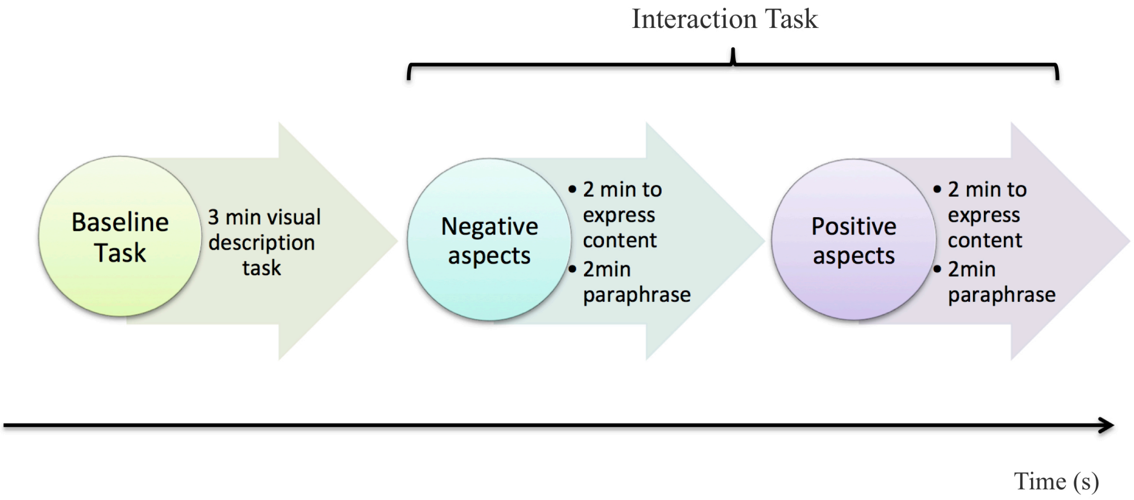


Figure I - 6 Schematic representation of the interaction task performed in laboratory. The selected sequence had in consideration ethical and methodological options of other marital interaction studies (e.g., Perrone-McGovern et al., 2014).

Afterwards, a team of independent coders segmented the entire interaction that lasted around 40 minutes into 20-second segments of negative and positive excerpts of the video. The most salient segments were then selected to construct the fMRI task. The final excerpts of the video were used to create the final fMRI experimental paradigm, which was developed in the software Psychopy v1.84 (Peirce et al., 2019).

The neuroimaging results from this fMRI paradigm reflected the individual´s brain activity during the performance of this ecological social cognitive task. This was the main focus of studies 2 and 3 of this dissertation. These studies were designed to build on the results of study 1 in which we looked at individual´s brain activity in a social brain network at rest. The full framework of this dissertation, and respective research questions, is described in the next section.

The present dissertation

Social neuroscience has been unveiling the neural correlates of social processes, such as the capacity to understand our own and others' internal states. Nonetheless, the neural correlates (and its distributed communication) of social cognition in the context of dyadic interaction, in particular romantic relationships, remain veiled. Addressing this gap in the literature, the main goal of the present dissertation was to investigate the neural correlates of social cognition in the context of a romantic relationship, using an advanced effective connectivity (EC) technique both with a resting state analysis and a more naturalistic fMRI paradigm. Chapters II to IV, correspond to three different research articles that compose the body of this empirical work; each describing different levels of analysis conducted throughout this doctoral dissertation.

Chapter II, which refers to study 1, focuses on the relationship between social cognition and the Default Mode Network (DMN). This well-known resting state network (Damoiseaux et al., 2006; Lu et al., 2012; Raichle et al., 2001; Raichle, 2015), has been consistently associated with a wide range of social cognition tasks involving perspective taking (Cox et al., 2012; Mars et al., 2012; Otti et al., 2010; Vemuri & Surampudi, 2015), affective and introspective processing (Amodio & Frith, 2006; Cavanna & Trimble, 2006; Forbes & Grafman, 2010; Mars et al., 2012; Saxe & Kanwisher, 2003) and self-other distinction (Sowden & Catmur, 2015; Steinbeis, 2015). However, despite the well-documented contribution of the DMN's key areas such as PCC, mPFC and bilateral IPL to social processing (e.g. Li et al., 2014; Mars et al., 2012; Schilbach et al., 2008; Schilbach et al., 2012) , evidence showing how the dynamic interplay within its nodes is related with social cognition is still lacking. Thus, in this chapter we used DCM to study how the functional integration within the DMN relates to affective and cognitive self-perceived social cognition abilities towards the romantic partner.

Chapter III, which refers to study 2, describes the use of a more ecological paradigm to characterize the socio-cognitive and socio-affective neural correlates involved in processing the internal states of self and others'. These "semi-independent" abilities, essential to our capacity to interact with others, seem to rely on similar brain mechanisms and processes. In this chapter we look at the neural basis of self and other processing in the context of a real social relationship. In particular, we used an fMRI paradigm adapted from a task resembling couples' natural interactions, where participants watched a series of videos-vignettes of his/her romantic partner expressing personal emotional contents and were asked to focus on his/her spouse´s experience or on his/her own experience. Thus the main goal of this study was to investigate the brain mechanisms involved in our ability to understand oneself and a close partner. Specifically, this study aimed to unravel which brain regions are commonly activated in both conditions and which ones are specifically involved in inferring about the self or about other.

In Chapter IV, we further explored how the flow of information changes within the social brain when we focus on our partner's internal states, in comparison to focus on the self. This study departed from a considerable gap in the literature, in respect to how the functional integration within the self and other common neural networks occurs, and how does it change during inferences about the self and others. To address this point, we used DCM to study the communication within a social cognition network activated during self and other processing in the experimental paradigm. In particular, this study explored which systematic changes in EC, among the bilateral TPJ, the PCC/prec and the bilateral MTG mediated the increase in rTPJ activation when inferring about other (rather than self).

Finally, in Chapter V we summarize and discuss the most relevant results of the empirical studies composing this doctoral dissertation, draw the most important conclusions, and depart from them to make some suggestions for the next steps in future research on the neural basis of social cognition.

CHAPTER II [1]

“Using resting-state DMN effective connectivity to characterize the neurofunctional architecture of empathy”

---

[1] This Chapter reproduces the publication:

Esmenio, S., Soares, J.M., Oliveira-Silva, P., Zeidman, P., Razi, A., Gonçalves, Ó.F., Friston, K., Coutinho, J., (2019). Using resting-state DMN effective connectivity to characterize the neurofunctional architecture of empathy. *Scientific reports* 9 (1), 2603

Abstract

Neuroimaging studies in social neuroscience have largely relied on functional connectivity (FC) methods to characterize the functional integration between different brain regions. However, these methods have limited utility in social-cognitive studies that aim to understand the directed information flow among brain areas that underlies complex psychological processes. In this study we combined functional and effective connectivity approaches to characterize the functional integration within the Default Mode Network (DMN) and its role in self-perceived empathy. Forty-two participants underwent a resting state fMRI scan and completed a questionnaire of dyadic empathy. Independent Component Analysis (ICA) showed that higher empathy scores were associated with an increased contribution of the medial prefrontal cortex (mPFC) to the DMN spatial mode. Dynamic causal modelling (DCM) combined with Canonical Variance Analysis (CVA) revealed that this association was mediated indirectly by the posterior cingulate cortex (PCC) via the right inferior parietal lobule (IPL). More specifically, in participants with higher scores in empathy, the PCC had a greater effect on bilateral IPL and the right IPL had a greater influence on mPFC. These results highlight the importance of using analytic approaches that address directed and hierarchical connectivity within networks, when studying complex psychological phenomena, such as empathy.

Keywords

Resting State; Default mode network; Functional connectivity; Effective connectivity; Dyadic empathy

Introduction

Neuroimaging studies have made important contributions to our understanding of the neural basis of several psychological processes (e.g. Lamm, Nausbaum, Meltzoff, & Decety, 2007; Otti et al., 2010; Singer, 2006). A variety of brain connectivity methods have been used to study the brain networks underlying these psychological processes, using functional magnetic resonance imaging (fMRI) data (see the comprehensive review in Soares et al. (2016)). Neuroimaging research in cognitive neuroscience, sometimes rely on functional connectivity (FC) methods (e.g., seed based analysis and independent component analysis - ICA) to describe the relationship between different brain regions (Cox et al., 2012; Ganis, Thompson, & Kosslyn, 2004; J. D. Greene, 2001).

However, the utility of functional connectivity methods is limited in social-cognitive studies, where the goal is to infer specific mental processes (e.g., theory of mind) from observed patterns of regional responses (e.g., in the temporal parietal junction and medial prefrontal areas). This inference is complicated by the fact that a single brain region is often recruited by a wide range of psychological processes and a given cognitive process may recruit more than one brain region (Poldrack, 2011; Price & Friston, 2002, 2005).

In fact, previous empirical evidence suggests that the more complex a given psychological process is, the more likely it is to engage the cooperation of several brain regions (Calhoun et al., 2001; Fan et al., 2011; Price & Friston, 2005; Tomasi & Volkow, 2012). This is true for complex psychological functions such as human empathy that comprises several processes; ranging from the ability to feel what the other person is feeling to the capacity to understand a given situation from the perspective of another person in order to anticipate his or her actions (Coutinho, Silva, & Decety, 2014). Due to the complexity of the construct of empathy the debate around its definition has been considerable, with some authors defining it in emotional terms (Batson, Fultz, & Schoenrade, 1987; Eisenberg & Strayer, 1987) and others defining it in cognitive terms (Hogan, 1969; Wispé, 1986).

One approach often used in the literature (Davis, 1994; Decety & Svetlova, 2012; Duan & Hill, 1996; Hoffman, 1984) to deal with this controversy is to adopt a multidimensional definition of empathy which includes both basic affective processes (e.g., emotional contagious and the ability to share affective states evoked by another individual), as well as more conceptual and cognitive dimensions (e.g., the ability to identify and understand the mental states of others). Similarly, in our work we define empathy as our ability to share, react to and understand the emotions of others. As we will see later, this was reflected in the empathy measure that we selected for this study, which assesses both emotional and cognitive dimensions of self-perceived empathy.

More specifically, in this study we focused on empathy expressed towards the romantic partner which is referred to as dyadic empathy (Long, 1990). Contrarily to general empathy that corresponds to the empathic tendencies in a general social context, not specific to a particular relationship, we were interested in looking at empathy expressed within the romantic relationship. Among the various dyadic interactions in which empathy has been studied, romantic relationships becomes an especially interesting context of study, since it is critically depending on feelings of compassion, support and empathic validation (Péloquin & Lafontaine, 2010). In fact, by allowing one partner to share and understand the other internal states, empathy is essential for stable and satisfactory couples relationships (Cohen & Strayer, 1996; Waldinger et al., 2004).

Extensive literature on empathy and the brain has consistently demonstrated that different neurobiological systems are involved in the various dimensions of empathy (for a review see Fan et al., 2011). For example, we know that experiencing another person´s feelings recruits emotional brain circuits comprised by anterior insula, parahippocampal gyrus, amygdala and anterior cingulate cortex (Bzdok et al., 2012; Fan et al., 2011), whereas our ability to cognitively understand other´s feelings and thoughts recruits the medial prefrontal cortex, temporal parietal junction and posterior cingulate cortex (Bzdok et al., 2012; Schurz et al., 2014).

The latter regions belong to the Default Mode Network (DMN) which is one of the better known resting state networks that comprises four key anatomical regions: the mPFC, the PCC and the left and right inferior parietal lobule (lIPL and rIPL) (Buckner, Andrews-Hanna, & Schacter, 2008; Greicius, Krasnow, Reiss, & Menon, 2003; Raichle et al., 2001; Raichle, 2015). This network whose regions seem to be recruited when we think about mental states, either ours and those of others, has been consistently associated with social functions in general (Mars et al., 2012; Leo Schilbach et al., 2008) and empathic abilities not only in humans (Mars et al., 2012; Otti et al., 2010; Vemuri & Surampudi, 2015) but also in rodents and primates . In fact, the DMN is of particular interest for the study of the neural basis of empathy, because it has been consistently associated with several psychological processes involved in it. The DMN has not only been associated with the understanding of another's pain through the inner representation of their affective states (Otti et al., 2010) but also with the ability to self-regulate the emotional arousal evoked by that vicarious experience in order to maintain an adequate distinction between our own and other´s psychological states.

Despite the well-documented role of the DMN for empathic processes, less is known about the dynamic interplay between its nodes and its importance for empathic processes. We consider that this is due to the fact that so far most studies looking at the involvement of the DMN in empathy used traditional functional connectivity approaches. In fact, as mentioned before, empathy is a very complex phenomenon and when analysing the brain responses underlying such high-level psychological processes, there is a need for analytic approaches that characterize the cortical hierarchies and directed causal relationships among brain areas within a network.

In contrast with functional connectivity (FC) approaches that are usually inferred on the basis of correlations among measurements of neuronal activity (i.e. which voxels display similar BOLD signal fluctuations over time), effective connectivity (EC) is defined as the directed influence that one brain region exerts over another (Friston, 1994, 2011)**,** supplementing FC in a complementary manner. While FC detects consistent spatiotemporal relationships between different brain regions, EC analyses considers how the information flow through these brain regions (Friston, 2009, 2011) providing insight into how the brain activity comes about.

In particular, dynamic causal modelling (DCM) infers effective connectivity (EC) between neuronal populations by combining dynamic models of neuronal states and detailed biophysical hemodynamic models (Friston, Harrison, & Penny, 2003). DCM supplements a forward model – of how cortical regions interact – with a hemodynamic model that transforms neuronal activity into measured response (e.g., blood oxygen level response) (Friston, Kahan, Biswal, & Razi, 2014; Razi, Kahan, Rees, & Friston, 2015; Sharaev, Zavyalova, Ushakov, Kartashov, & Velichkovsky, 2016). DCM has been used extensively to study task-related brain responses (Allen et al., 2010; Kiran, Meier, Kapse, & Glynn, 2015; Osnes, Hugdahl, & Specht, 2011; Sonty et al., 2007; Yu, Chen, Liu, & Zhou, 2013); however, recent DCM methods now also support inference on connectivity during resting-state – by modelling endogenous neural fluctuations (Friston et al., 2014; Razi et al., 2015): this is referred to as spectral DCM. DCM of effective connectivity at rest has been applied to clinical phenomena like melancholia (Hyett, Breakspear, Friston, Guo, & Parker, 2015); schizophrenia (e.g. Bastos-Leite et al., 2015; Cui et al., 2015) and smoking addiction (Tang et al., 2016).

The use of DCM in this context may also be beneficial to the understanding of normal psychological processes, such as empathy, that entail distributed processing over large-scale networks. Thus, in this study we used spectral DCM (Friston et al., 2014; Razi et al., 2015), to analyze the effective connectivity within the DMN – and to quantify how the functional architecture of this brain network relates to self-perceived empathy. Our reasoning was that by understanding how the information flow among DMN regions is related with differences in empathy abilities, we will better understand the role of each node in empathy (e.g. Menon & Uddin, 2010). For this purpose, we used both a functional (by means of an ICA analysis) and effective connectivity (by means of a DCM analysis) to analyze the relationship between the DMN functional architecture and self-perceived empathy. In summary, our hypothesis was that variations in the self-perceived empathy would be underwritten by systematic differences in effective connectivity among the nodes of the default mode network – and that these differences in directed coupling mediate differences in functional connectivity.

Methods

Participants

Forty-two individuals, comprising 21 heterosexual couples that reported being in committed monogamous romantic relationship for at least one year, participated in this study. Participants' ages ranged from 23 to 40 years old (M = 31.17, SD = 4.748; for males: M = 32.13, SD = 4.893, for females: M = 30.22, SD = 4.502). Prior to any procedure, all participants were screened on the telephone to assess inclusion and exclusion criteria. Exclusion criteria were as follows: (1) the presence of any dementia and/or diagnosed neuropsychiatric and/or neurodegenerative disorder; (2) dependency or abuse in the past year of alcohol and/or drugs; (3) incapacity and/or inability to attend the MRI session (e.g., metallic implants; pregnancy); (4) age bellow 20 or above 50 years. All participants were Caucasian and right-handed. The study goals and procedures were explained and all the participants gave informed written consent. All the procedures were approved by the Institutional Review Board of University of Minho and the study was conducted in accordance with the principles expressed in the Declaration of Helsinki.

Empathy Measures

To assess dyadic empathy, participants completed the Portuguese version of the Interpersonal Reactivity Index for Couples (IRIC)(Coutinho et al., 2016; Péloquin & Lafontaine, 2010). The IRIC is a modified version of the Interpersonal Reactivity Index (IRI) that assesses cognitive and emotional empathy in the context of intimate relationships. It is composed by 13-items that assess empathy toward the partner specifically (e.g.: *"I* sometimes try to understand my partner better by imagining how things look from his/her perspective"). The cognitive aspects of empathy are measured by the Dyadic Perspective-Taking (DPT) subscale, which evaluates individuals' ability to adopt the psychological point of view of their partner. The emotional aspects are measured by the Dyadic Empathic Concern (DEC) subscale, which evaluates individuals' feelings of concern for their partner. Each questionnaire item is measured on a 5-point Likert-type scale, with responses ranging from: (0) "Does not describe me well" to (4) "Describes me very well". Higher scores in a scale represent greater perspective-taking or empathic concern. The IRIC total score can vary between 0 and 52, with a higher score indicating higher perceived empathy within the couple's relationship. The score of dyadic perspective taking can vary between 0 and 24 and the score of dyadic empathic concern can vary between 0 and 28. The Portuguese version of IRIC showed adequate internal consistency values (Cronbach's α =0,82) and good levels of external validity with a dyadic adjustment scale (DAS)(Coutinho et al., 2016; Péloquin & Lafontaine, 2010).

The administration of the IRIC occurred before the neuroimaging acquisition and took place in the lab using a paper and pencil version. The IRIC was administered anonymously and independently to each element of the couple, that is, each spouse completed the questionnaire individually and was asked not to exchange impressions on items or their answers with the partner. A member of the research team was present in the room to ensure the participants had no doubts about the IRIC. In order to control for social desirability effects and ensure the total confidentiality of the answers a code was assigned to each participant and to his/her respective questionnaire, which, after completed, was introduced in a sealed envelope. The values obtained in our sample for the scores are comparable to those reported with other samples (Coutinho et al., 2016; Péloquin & Lafontaine, 2010) :DPT, M = 16.37, SD = 3,83 and DEC, M = 23,72, SD = 2.951.

MRI acquisitions

MRI data were acquired with a clinically approved 3Tesla MRI scanner (Siemens Magnetom Tim Trio, Erlangen, Germany). A 7-min resting-state functional acquisition (210 volumes) was obtained for each participant using a sensitive echo-planar imaging (EPI) BOLD sequence with the following imaging parameters: repetition time (TR) = 2000 ms; echo time (TE) = 29 ms; flip angle = 90◦; field of view (FoV) = 1554 mm; matrix size = 64 × 64; pixel size = 3 × 3 mm^2, slice thickness = 3 mm and 39 slices. Additionally, one structural scan (192 sagittal slices, repetition time (TR) =2000 ms; echo time (TE) = 2.33 s, flip angle = 7°, slice thickness = 0.8 mm, slice gap = 0 mm, pixel size = 0.8 × 0.8 mm2, field of view (FoV) = 256 mm) was also acquired. Participants were instructed to keep their eyes closed and to remain awake but relaxed, doing nothing during the acquisition and remaining as motionless as possible. None of the participants fell asleep during the acquisition.

Functional connectivity pre-processing and analysis

Image pre-processing and spatial independent component analysis (ICA) was performed with tools provided with the FMRIB Software Library (FSL v5.09; http//fsl.fmrib.ox.ac.uk/fsl/) (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012; Smith et al., 2004). Preprocessing steps included: removal of the first five volumes (10 s) to ensure signal stabilization and allow the participants to adjust to the scanner noise; slice-timing correction using the first slice as reference; motion correction (using rigid body alignment of each volume to the mean image of the acquisition using MCFLIRT) & motion scrubbing (volumes in which FD > 0.5 and DVARS > 0.5% change in the BOLD signal were "scrubbed," or removed entirely from the data (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012); non-linear normalization to the MNI standard space (using the structural T1 normalization matrix); regression of motion parameters, mean WM and CSF signals (generated by segmenting each individual's structural image using FAST, thresholded to ensure 80% tissue type probability); band-pass temporal filtering (0.01–0.08 Hz) and spatial smoothing (8 mm full width at half-maximum Gaussian kernel) (Soares et al., 2016). Before any data processing and analysis, all acquisitions were visually inspected to confirm that they were not affected by undue head motion and that participants had no brain lesions. On average, 6 volumes were removed per participants (scrubbing correction) and three participants were excluded due to the presence of head motion of greater than 3 mm in translation or 1.5° in rotation.

Next, the spatial ICA was performed for each participant, but across all participants. Briefly, spatial ICA analysis is a multivariate data-driven approach that uses temporal correlations to compute spatial maps of functionally connected regions, while maximizing the independence between them (Power et al., 2012). To perform the spatial ICA, first we used MELODIC to search for common spatial patterns among subjects and automatically estimate the number of independent components to extract. After performing the group ICA, dual regression was applied to recover each subject's version of the group components. Pearson correlation coefficients were then calculated between the time-series of each pair of components ROIs, resulting in one correlation matrix per subject, which were then transformed to Z-score matrices by the application of Fisher’s transform. The final 21 components identified by ICA were visually inspected, sorted, and associated with the resting state functional networks from (Soares et al., 2016). This revealed 11 independent resting state networks with typical spatial patterns of functional connectivity. The component that corresponded to the DMN was identified for the final sample (N=39). The resulting independent component or spatial mode was used as a summary of functional connectivity (implicit in the ICA).

Our initial analysis comprised a one-sample t-test in SPM12 (Wellcome Department of Cognitive Neurology, London, UK) to identify regions within the default mode that were conserved over subjects. A second analysis was then performed to test for a relationship between dyadic empathy – as assessed by the IRIC and the expression of the DMN. Statistical parametric mapping was computed at the between subject level using subject specific DMN maps as the response variable and IRIC’s total score as the explanatory variable in a general linear model. This model included age and gender as nuisance variables. Results were considered significant at $p < 0.05$, corrected for multiple comparisons using the Monte Carlo correction. The correction was determined for clusters with a height threshold of $p \leq 0.01$ and a size threshold of 116 contiguous voxels. A combination of visual inspection and Anatomical Automatic Labeling atlas (AAL) was used for anatomical labelling (Tzourio-Mazoyer et al., 2002).

Effective connectivity pre-processing and analysis

In a second data analysis stream, preprocessing and subsequent dynamic causal modelling were performed using the Statistical Parametric Mapping software (SPM12; Wellcome Department of Cognitive Neurology, London, UK). The preprocessing steps included: removal of the first five volumes; slice-timing correction; motion correction (re-aligned to the mean image); normalization of the functional acquisition to the MNI standard space through the sequential application of a rigid body transformation and the nonlinear warp resultant of previous nonlinear registration of the structural scan to the MNI T1 template (Ashburner & Friston, 2000); regression of motion parameters, mean WM and CSF signals (generated by segmenting each individual's structural image using SPM, thresholded to ensure 80% tissue type probability) and smoothing using an 8 mm full width at half-maximum Gaussian kernel. After processing, images were visually inspected to ensure that they had not any undue head motion and that participants had no brain lesions. As above, three participants were excluded as they exceed a head motion higher than 3 mm (translation) or 1.5° (rotation).

Spectral DCM (spDCM) was used to analyze the resting state fMRI data. In brief, spDCM analysis involves a specification of a plausible network model, which then enables the estimation of the model parameters that quantify effective connectivity and regionally specific haemodynamic variables (Friston et al., 2014; Razi et al., 2015). Model specification comprised the selection of the regions of interest and definition of the model space in terms of connectivity between regions. Based on the previous literature on DMN neuroanatomy (Buckner et al., 2008; Greicius et al., 2003; Mars et al., 2012; Raichle et al., 2001; Raichle, 2015), we identified four ROIs as key DMN nodes; including the mPFC, PCC, left and right inferior parietal lobule.

The MNI coordinates for these nodes were identified for each participant using previously described ICA analysis. The regions of interest were masked by a (8 mm radius) sphere centered at the participant-specific coordinates. Since there is no previous literature on the relationship between the information flow within DMN and empathy, we adopted and exploratory approach, starting with a fully connected model, i.e. a model in which all the ROIs communicate with each other. A fully connected model was constructed for each subject, containing a total of 16 connectivity parameters: 12 connection between regions and 4 recurrent self-connections (see Figure 2 c). Having specified the full DCM, it was inverted for each participant. Furthermore, Bayesian model reduction (BMR) (Friston & Penny, 2011;

Rosa, Friston, & Penny, 2012) was used was used to find for each participant the best model (i.e. the relevant connections) to explain the data. BMR selected the model with the highest posterior probability from every possible design nested within the previously defined model (i.e. fully connected model).

Afterwards, to ensure convergence, cross spectra density (CSD) data fit was inspected visually. One participant was excluded due to a poor data fit. Finally, to study the relationship between DMN's DCM effective connectivity estimates and empathy scores , we performed canonical variate analysis (CVA) (Darlington, Sharon, & Walberg, 1973). In the CVA, participant's DCM data was combined with their scores in each subscale of IRIC's questionnaire (i.e. perspective taking and empathic concern scales). This way it was possible to understand the relative influence of each IRIC's subscale in the changes in DMN's connectivity. The canonical vectors were considered significant at $p < 0.05$.

Results

Functional connectivity results

The spatial modes (i.e., independent components) of FC in the DMN during resting state were identified at the group level and four main regions were observed; namely, the PCC and precuneus (PCu) (x= 2; y= -60; z= 30; Z>8, 9767 voxels), the bilateral IPLs (x= -40; y= -70; z= 40; Z>8, 2339 voxels and x= 44; y= -64; z= 32; Z= 7.74, 1746 voxels) and the mPFC (x= -6; y= 52; z= 8; Z= 7.60, 1941 voxels) (Figure 1). At rest, increased IRIC total score was associated with an increased expression of the DMN in the mPFC, primarily in the frontal pole (peak MNI x= 10; y= 50; z= -10; Z=2.92, 117 voxels) (Fig II - 1).

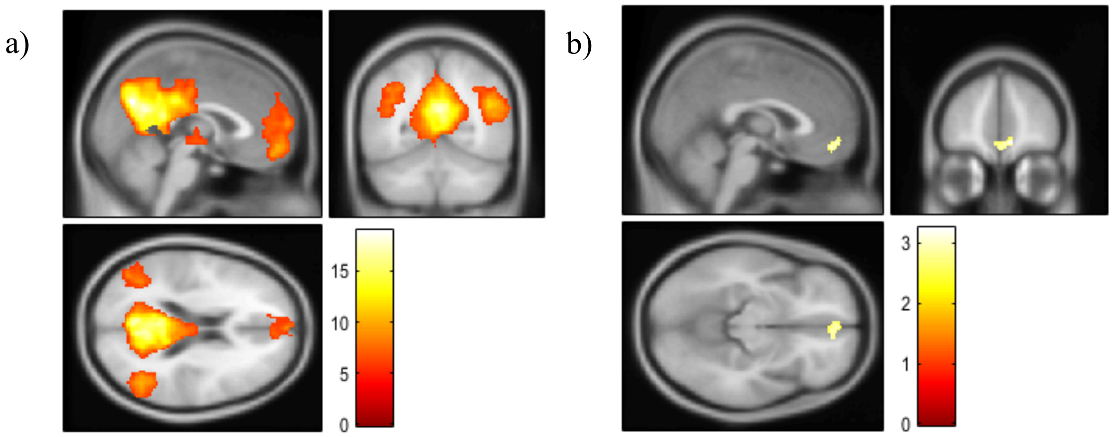


Figure II -1 a) Consistent DMN expression identified in resting state over subjects. b) Correlation between IRIC total score and DMN functional connectivity

Effective connectivity results

Canonical variate analysis of the multivariate relationship between DCM effective connectivity estimates and dyadic empathy scores showed a significant canonical correlation for the first canonical vector (p = 0.01598, chi-squared statistic (X^2) = 33.1996). This vector showed that higher IRIC total scores (M = 40.36, SD = 4.50) were associated with an overall decrease in connectivity between regions and regional self-inhibition (Figure II - 2 a).

To understand the importance of these results, an important distinction needs to be made between extrinsic or direct connections (i.e., connections that reflect the EC between regions - represented in blue in Figure 2 a) and intrinsic or self-inhibition connections (i.e. connections that reflect how susceptible a node is to extrinsic afferent - represented in green in Figure 2 - 2 a). In terms of extrinsic connections, there was a general decrease of the influence of all DMN nodes on each other, except from: PCC to rIPL (4); rIPL to PCC (13); PCC to lIPL (3) and rIPLto mPFC (14) (Figure II - 2 a and d). In other words, there was a general decrease in effective connectivity accompanied with a selective increase in the influence of PCC on bilateral IPL, and a right hemisphere increase in the influence of IPL on the medial prefrontal cortex.

Regarding the intrinsic connections, the results show lower self- inhibition in all nodes with higher empathy scores; (Figure 2 - 2a and d) specially in PCC(6) and lIPC (11) (Figure 2 - 2 d). In this case, a smaller self-inhibition in a node represents an increase in sensitivity to extrinsic afferents, i.e. the node becomes more excitable, taking less time to respond to the influence of other nodes.

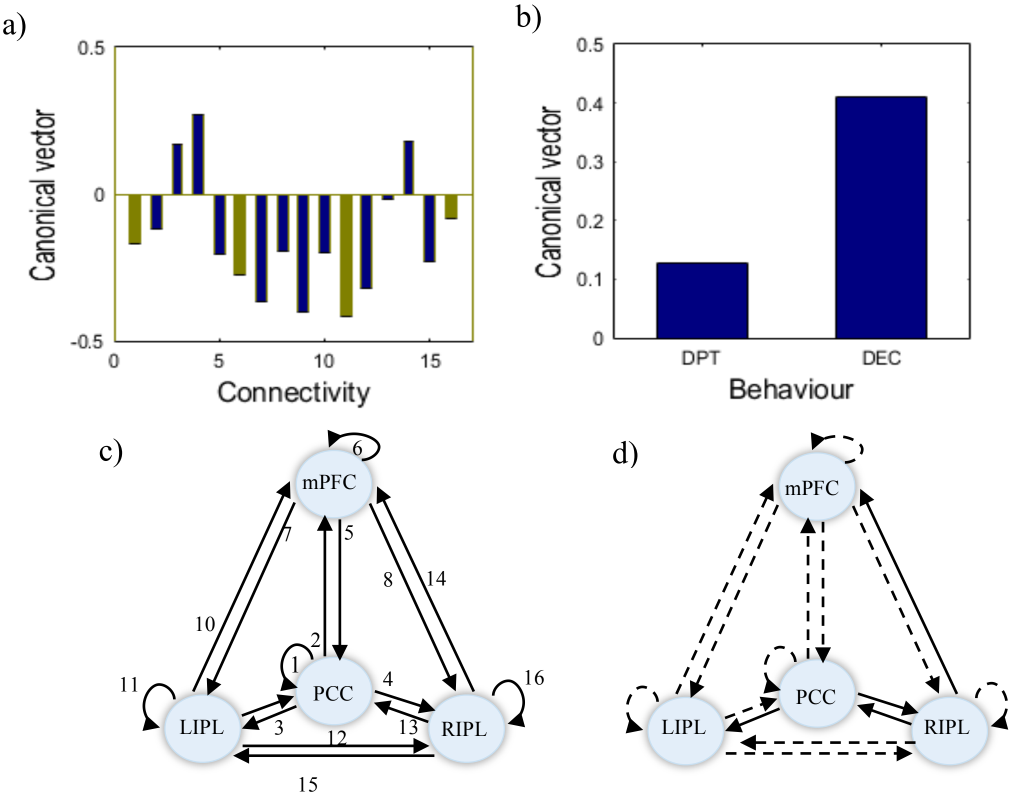


Figure II - 2 a) Canonical Vector containing weight (contribution) assigned by CVA to each DCM effective connectivity parameter. Blue connections correspond to extrinsic connections and green to intrinsic connections. b) Canonical Vector containing weight (contribution) assigned by CVA to each empathy scale score: perspective taking (PT) and empathic concern (EC). c) Showing the index number of each effective connection in (a) d) weight of each connection, where bold lines represent an increase and dashed lines a decrease with empathy.

In addition, this canonical vector also highlights that empathic concern in comparison with perspective taking, has a greater contribution to the canonical correlation; i.e., a greater predictive validity in terms of EC within the DMN (Figure II - 2 b).

Discussion

The current study examined the extent to which individual differences in self-reports of dyadic empathy were reflected in functional integration within the DMN. Differences in DMN connectivity were addressed both in terms of functional and effective connectivity among different brain areas, by means of ICA and DCM respectively. Our results showed significant correlations between DMN's functional architecture and self-reported dyadic empathy.

In terms of functional connectivity, ICA showed that higher scores in dyadic empathy were associated with a significant increase in the expression of the mPFC; more specifically, in the frontal pole. The mPFC has been linked with our capacity to reflect on the mental state of others (Amodio & Frith, 2006; Gilbert et al., 2006), and the frontal pole has been shown to play a crucial role in the functional integration between cognition and emotion; particularly, in the case of adaptive empathic responses (Carrington & Bailey, 2009; Decety & Jackson, 2004; Decety & Svetlova, 2012). In line with our results, previous literature using ICA showed that when compared to a medium-empathy group, low-empathizers display lower FC of the mPFC within the DMN (Kim et al, 2017). This evidence suggests a relationship between empathy and the functional connectivity of the mPFC within the DMN. Between-subject differences in functional connectivity underlying this result were clarified, mechanistically, in terms of differences in effective connectivity involving not just afferents to the medial prefrontal cortex but other key differences throughout the DMN.

In terms of directed coupling within the DMN, spectral DCM (Friston et al., 2003, 2014; Razi et al., 2015) showed a clear and consistent pattern. Increase in empathy scores was associated with a general decrease in EC between DMN nodes that was accompanied with a universal disinhibition in all nodes. In other words, not only did nodes became more self-sufficient but, simultaneously, showed a critical slowing with decreased self-inhibition. This effect was accompanied by an increase in PCC´s influence on bilateral inferior parietal cortices. Essentially, as self-perceived empathy scores increases, the PCC seems to dominate interactions within the DMN. These results are consistent with previous evidence that speaks to the PCC´s role as a connector node for the functional integration within and between brain networks (Deshpande, Santhanam, & Hu, 2011; Hagmann et al., 2008; Heuvel & Sporns, 2011).

It is interesting that the particular connections from the PCC that increased with empathy over subjects were in the descending direction (i.e., they were PCC efferents as opposed to ascending afferents). This may speak to the fundamental role of top-down predictions about one's own bodily state when trying to infer the intentions of others (Kilner, Friston, & Frith, 2007); particularly in the context of predictive coding formulations of empathy (Brown & Brüne, 2012; Gu et al., 2015; Krahé, Springer, Weinman, & Fotopoulou, 2013). Additionally, we found that the PCC exerted an excitatory coupling on both IPLs that propagated vicariously to the mPFC, via the rIPL. In other words, participants with higher self-perceived empathy presented with a lateralization in the DMN, with the right IPL acting as the mediator between PCC and mPFC.

This lateralization – through the rIPL – within the DMN is consistent with lesion studies which suggest that empathic abilities are compromised in both patients with prefrontal lesions and patients with lesions on the right parietal cortex (Shamay-Tsoory, Tomer, Goldsher, Berger, & Aharon-Peretz, 2004). In fact, when lesions involved the left parietal cortex, patients showed no impairment on empathic ability. This highlights the specific contribution of the right IPL – and mPFC – in empathic processes.

Furthermore, previous research assessing the functional integration in the DMN showed a clear rightward asymmetry of extrinsic connectivity (Di & Biswal, 2014). In addition, this study showed that the right IPL sends information both to the mPFC and the PCC, and the mPFC sends information to the PCC. Therefore, it seems that the higher self-perceived empathy is associated with a decrease in the communication between the mPFC and the PCC and with an inversion of the information flow between rIPL and PCC.

In summary, in agreement with the consistent relationship found in the literature between the DMN and human empathy (Mars et al., 2012; Otti et al., 2010; Vemuri & Surampudi, 2015), our results show a pattern of intrinsic connectivity changes in this network associated with differences in dyadic empathy measures. Furthermore, systematic differences in information flow – among the DMN regions – showed that in individuals with higher dyadic empathy, the PCC seems to assume a central function within the DMN, by influencing, directly, the bilateral parietal regions and, indirectly, the mPFC through the right IPL. Although in the functional connectivity analysis the mPFC was the node associated with higher self-perceived empathy, the results from the DCM analysis suggest that this increase results from a shift in coupling between PCC and rIPL. These results highlight the importance of applying effective connectivity analysis, such as DCM, when trying to characterize more complex psychological phenomena, such as empathy.

Finally, the results of the CVA analysis (Darlington et al., 1973) include the relative contribution of each of the IRIC scales (i.e., empathic concern and perspective taking) to the previously mentioned differences found in DMN connectivity. Results showed that while both subscales contributed positively to the observed changes in connectivity within the

DMN, the empathic concern subscale presented a significantly higher contribution to those changes. In previous studies, Schilbach et al. (2008) reports that DMN acts as a physiological "baseline" of the human brain that is linked to our predisposition for social cognition. Similarly, our results suggest the DMN as a network that is linked to more 'affective–perceptual' forms of empathy, hence empathic concern.

Although our results are compelling, the current study is not free from the limitations inherent to all the resting state studies; such as the difficulty with controlling behavior related confounds such as sleep and the unconstrained nature of the "task". In fact, test–retest reliability of resting state connectivity measures typically range between moderate to good even with optimal processing (Braun et al., 2012; Cao et al., 2014), which may be due to the absence of cognitive constraints. Future studies could try to monitor the cognitive processes at rest by using some measures such as mind wandering questionnaires – or indeed a carefully designed empathy task to look at condition-specific changes in effective connectivity. Furthermore, it would be interesting to apply advanced (second level) data analysis method in combination with DCM (e.g. parametric empirical Bayes – PEB), to assess the specific contribution of each self-report subscale (e.g. empathic concern) on each specific connection (e.g. PCC to rIPL). Finally, it could be interesting to examine the existence of similar connectivity patterns in other resting state networks that include regions associated with empathy response, namely the Salience Network that includes both the anterior insula & dorsal ACC (Toyomaki & Murohashi, 2013; Uddin, 2017).

Acknowledgements and Author Contributions

This study was funded by BIAL Foundation (Grant number 87/12); by the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Education and Science through national funds and co-financed by FEDER through COMPETE2020 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653); by the postdoctoral scholarship UMINHO/BPD/18/2017 and by the FCT Doctoral scholarship (PD/BD/105963/2014).

J.C. and O.F. G. designed the research, J.C. and P.O-S. collected data for the experiments. S.O., J.M.S., P.Z., A.R. and K.F. analyzed the collected data. J.C., J.M.S. and S.O. wrote the paper and all authors reviewed and approved the final draft.

CHAPTER III[2]

“Brain circuits involved in understanding our own and other´s internal states in the context of Romantic Relationships”

---

[2] This Chapter reproduces the publication:

Esmenio, S., Soares, J.M., Oliveira-Silva, P., Gonçalves, Ó.F., Decety, J., Coutinho, J., (2019). Brain circuits involved in understanding our own and other´s internal states in the context of Romantic Relationships. *Social Neuroscience,* DOI: 10.1080/17470919.2019.1586758

Abstract

Social interactions require the capacity to understand both our and other’s internal states. These semi-independent skills, the ability to understand oneself and others, seem to rely on the same type of representations and recruit similar brain areas. In this study, we looked at the neural basis of self and other processing in the context of an interaction with a significant other. Fourty-two participants in a monogamous relationship completed an fMRI task in which they watched a set of video-vignettes of his/her romantic partner expressing emotional contents. Participants were asked to elaborate on his/her spouse´s experience (other condition) and on his/her own experience when watching the video-vignettes (self-condition). The results showed a significant overlap in the brain activation for both conditions (e.g. anterior insula, posterior cingulate/ precuneus, inferior frontal gyrus, inferior parietal lobule). In addition, the self-condition recruited brain areas associated with interoceptive processing and affect sharing (e.g., posterior insula), whereas the other-condition engaged brain areas involved in the cognitive representation of another’s internal states and self-other distinction (e.g., fusiform, supramarginal gyrus, angular gyrus and temporoparietal junction).

Keywords: Social cognition; Theory of mind; Self / other processing; Romantic relationships;

Introduction

The ability to make sense of what we ourselves and others feel or think is at the center of our capacity to interact with others. There is a general consensus, among social neuroscience researchers, that this capacity to understand internal states, either ours or those of others, recruits two main pathways: a socio- affective route (de Vignemont & Singer, 2006), and a socio-cognitive route (C. Frith & Frith, 2005). The socio-affective pathway corresponds to embodied, visceral or bottom up dimensions of social processing (Jean Decety & Lamm, 2006; Singer & Lamm, 2009), while the socio-cognitive pathway refers to conceptual or top down processes that involves higher cognitive functions (Carrington & Bailey, 2009; C. Frith & Frith, 2005).

Research shows that these two pathways engage independent brain networks (Kanske, Böckler, Trautwein, & Singer, 2015). The socio-affective route recruits areas such as the anterior cingulate cortex (ACC), anterior midcingulate cortex (aMCC) extending into the supplementary motor area (SMA), anterior insula (AI), inferior frontal gyrus (IFG), posterior cingulate (PCC), anterior thalamus, the middle temporal gyrus (MTG),midbrain and inferior parietal lobe (IPL) -including the supramarginal gyrus (SMG) and angular gyrus. (Bzdok et al., 2012; Fan et al., 2011). On the other hand, the socio-cognitive pathway which are often termed as Theory of Mind (TOM), or mentalizing recruits a network of brain regions that includes the posterior/ventral temporoparietal junction (TPJ), precuneus, MTG, superior temporal sulcus (STS), temporal poles and medial prefrontal cortex (mPFC) (Bzdok et al., 2012; Schurz et al., 2014).

Although both socio-affective and socio-cognitive pathways have been extensively investigated, research has only recently start looking at how these two pathways interact when processing our own and other´s internal states in situations in real social life (Lamm, Decety, & Singer, 2010; Zaki, Weber, Bolger, & Ochsner, 2009). In fact, the majority of the experimental studies in social cognition use paradigms with fictional targets (e.g., Singer et al., 2006) or relatively artificial situations, with very few studies looking at the brain´s response in real interpersonal interactions. This is important because in naturally occurring social encounters there are several factors that influence the neural response to other´s emotions such as the affective proximity with the target, familiarity (Cialdini, Brown, Lewis, Luce, & Neuberg, 1997), and how much protection or care (Batson et al., 2005) the target needs. Another aspect that characterizes our most significant interpersonal interactions such as parent-child or couples' interaction, is that we often have to empathize with emotional contents that are directed towards us and thus often have to alternate between the focus on our internal states and those of the other

Research in social neuroscience suggests that the understanding of other's feelings and mental states rely on the same type of representations necessary to understand our own internal states (Lamm, Bukowski, & Silani, 2016). Accordingly, previous studies have found that similar brain areas are recruited when processing self or other internal states in both affective (e.g. empathy for pain) (Lamm et al., 2010; Singer et al., 2004), as well as cognitive experimental paradigms (e.g. reasoning about oneself/ mentalizing) (Lombardo et al., 2009; Mitchell et al., 2006). In agreement with these findings, Dimaggio, Lysaker, Carcione, Nicolò, & Semerari (2008) suggested that the ability to understand oneself and others are “semi-independent skills". Indeed, recent evidence has also found that enhancing the ability to understand our own states increases the ability to understand those of another (Böckler et al., 2017). This may be particularly relevant for the context of romantic relationships, if we consider evidence from simulation theories of social cognition suggesting that the closer the other is to oneself, the more likely we are to base inferences about them on the knowledge about oneself (Adolphs, 2002; Vittorio Gallese & Goldman, 1998).

However, as mentioned before, so far, the attempts to investigate the interdependence of self and other processing in the context of a meaningful relationship with a significant other, in particular in romantic relationships, are still scarce. Hence, the present study investigates the brain mechanisms involved in our ability to understand oneself and a romantic partner.

Indeed, romantic relationships are the central relationship for most adults, having an important duration in human´s life cycle (Robles & Kiecolt-Glaser, 2003). They constitute a rich source of emotional interchange (Gottman & Levenson, 1988), highly dependent on feelings of understanding, support and empathic validation (Péloquin & Lafontaine, 2010; Waldinger et al., 2004). In addition, our choice for focusing on romantic relationships was justified by the well-documented negative impact of marital conflict on the physical and mental health of each partner (e.g., Beach, Fincham, & Katz, 1998; Coyne et al., 2001; Greene & Griffin, 1998), namely with important alterations in cardiovascular, endocrine and immune function (Mayne et al., 1997). Finally, marital conflict has important social consequences not only for those directly involved in the conflict, but also for other family members, particularly children (Cummings & Davies, 2002).

In order to address our research question we built an ecological functional magnetic resonance imaging (fMRI) task in which participants were instructed to watch a set of emotional video vignettes of their respective partner and was asked to elaborate on his/her partner´s experience (other condition) or on his/her own experience (self-condition). Therefore, the main aim of this study is to determine the common and specific brain pathways involved in both the processing of the individual´s own internal states as well as in the processing of a significant other's internal states.

Regarding the common pathways, we predict a significant overlap between the self and other brain activations. This overlap should include socio-affective areas, such as ACC, aMCC, SMA, AI, PCC, anterior thalamus, IFG, IPL and midbrain (Bzdok et al., 2012; Fan et al., 2011), as well as socio- cognitive areas such as the TPJ, precuneus, STS, temporal poles and mPFC (Bzdok et al., 2012; Schurz et al., 2014). On the other hand, the capacity to understand one's internal states and to understand those of another should not be reducible to one another. Accordingly, we hypothesize that the self condition will also recruit brain areas associated with sensory and interoceptive processing, such as the posterior insula, somatosensory cortex and the SMA (Chang, Yarkoni, Khaw, & Sanfey, 2012; Craig, 2003; Khalsa et al., 2017; Singer et al., 2004); while the other condition should engage to a further extent brain areas involved in mentalizing/ TOM, including areas of perceptual processing, such as the inferior occipital gyrus, fusiform gyrus, superior temporal gyrus (Adolphs, 2002), and regions responsible for self-other distinction, namely the rTPJ and rSMG (Lamm et al., 2016; Steinbeis, 2015).

Method

Participants

Forty-two participants (N=42) in a committed monogamous romantic relationship for at least one year participated in this study. Prior to any procedure, all participants were screened on the telephone to assess inclusion and exclusion criteria. The exclusion criteria included: (1) the presence of any dementia and/or diagnosed neuropsychiatric and/or neurodegenerative disorder; (2) dependency or abuse of alcohol and/or drugs in the past year; (3) inability to attend the MRI session (e.g., metallic implants; pregnancy); (4) age below 20 or above 50 years.

All the participants were Caucasian, right-handed and the age of the participants ranged from 23 to 40 years old (M = 31.17, SD = 4.748; for men: M = 32.13, SD = 4.893, for women: M = 30.22, SD = 4.502). Regarding the characteristics of the relationship: (1) 30,4% of the couples were married; (2) 39,1% were living together; (3) 30,4% were in dating relationships; (4) 34.8% of the couples had children. In the final sample the duration of the relationships ranged from 1 to 15 years (M= 7.78; SD=4.17). Before the implementation of the study all the goals and procedures were explained to the participants that provided an informed written consent. All the procedures were approved by the University Institutional Review Board and complied with the principles expressed in the Declaration of Helsinki (with the amendment of Tokyo 1975, Venice1983, HongKong 1989, Somerset West1996, Edinburgh 2000).

Experimental Task

In this task each participant watched a series of videos-vignettes of his/her romantic partner expressing personal emotional contents, and was asked to focus on his/her spouse´s experience (other condition) or on his/her own experience (self condition) while listening to the content expressed in the video. The video-vignettes used in this task were extracted from a previously recorded real interaction task performed in the lab, where the couple communicated positive and problematic topics in their relationship. In this interaction task, participants were asked to disclose, one at a time, things that they would either like or dislike concerning their partner. The negative and positive aspects were selected by the spouses from a list of possible topics that was provided to them before the interaction. Participants could choose other relevant topics to discuss. Examples of negative topics included things that made them feel upset or bothered, things that they would like to change or things they didn't like about their partner. Examples of positive topics included things that made them feel happy, things they truly admire about their partner, things that they would like to keep as it is.

This interaction was videotaped and a team of independent coders segmented the entire videotaped interaction task and selected the most salient 20-second negative and positive excerpts of the video. These negative and positive video vignettes were then used to construct a functional magnetic resonance imaging (fMRI) paradigm. Neutral videos were extracted from the Emotional Movie Database (Carvalho et al., 2012). The specifics of this interaction task are described in further detail in (Coutinho et al., 2017; Coutinho et al., 2018).

The task consisted of two blocks (self or other condition), where each block contained 22 trials comprising the three different conditions: positive communications (8 trials), negative communications (8 trials) and neutral videos (6 trials). Each trial consisted of: (1) fixation cross (5 seconds); (2) instructions in accordance with the present block (e.g. “In the next movie focus on how your partner is feeling.”) (3 seconds); (3) video-vignette (20 seconds); (4) behavioural response (4 seconds). An example of an emotional and a neutral trial is displayed in Figure 1.Regarding the behavioural response, participants were required to choose among one of three possible responses dependent on the emotional impact of the vignette: (1) “Bad” for any kind of discomfort, negative state or emotion - e.g. distress, feeling rejection, incomprehension, abandonment; (2) "Neutral" in the absence of any positive or negative state or emotion; or (3) “Good” for any kind of positive state or emotion – e.g. happiness, ease, acceptance, love, connection or comprehension.

Stimuli were displayed in a pseudo-randomized order that remained the same across blocks. Finally, blocks were displayed in a randomized order across participants. The total duration of the task was 1364s (24min).

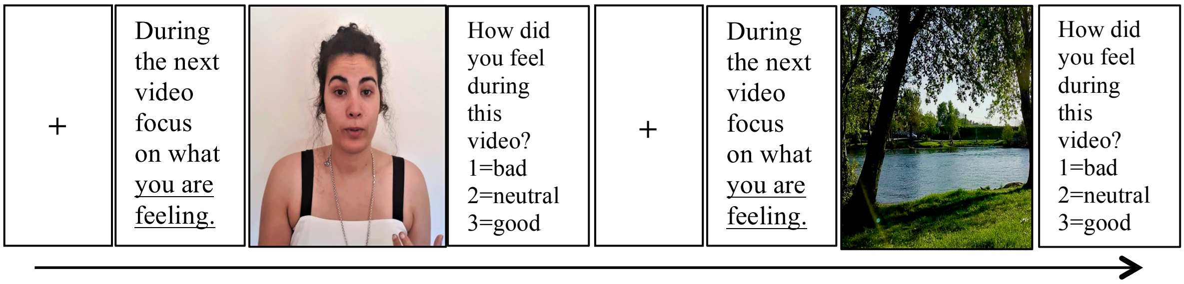


Figure III-1 Scheme of an emotional and a neutral trial in the self-condition.

Neuroimage acquisition

Structural and functional images were acquired with a clinical approved 3T MRI scanner (Siemens Magnetom Tim Trio, Erlangen, German). The imaging session comprised one structural scan (192 sagittal slices, repetition time (TR) =2000 ms; echo time (TE) = 2.33 s, flip angle = 7°, slice thickness = 0.8 mm, slice gap = 0 mm, pixel size = 0.8 × 0.8 mm^2, field of view = 256 mm) and one functional blood oxygen level dependent (BOLD) sensitive echo-planar imaging (EPI) sequence (39 axial slices; TR= =2000 ms; echo time (TE) = 29 ms, flip angle = 90◦, matrix size = 64 × 64, slice thickness = 3 mm, pixel size = 3 × 3 mm^2, field of view = 222x222 mm) were collected for each participant. The synchronization between the paradigm and the acquisition were guaranteed for each TR.

Data preprocessing and analysis

The preprocessing and subsequent data analyses were performed using the Statistical Parametric Mapping software (SPM12; Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk). The preprocessing steps comprised: (1) slice-timing correction, (2) motion correction to correct for movement artifacts and related susceptibility artifacts through the re-alignment to the mean image; (3) coregistration of the anatomical and the functional images through a rigid-body registration of the mean functional image to the structural scan; (4) normalization of the functional acquisition to the Montreal Neurological Institute (MNI) standard space (Ashburner & Friston, 1999) through the sequential application of a rigid body transformation and the nonlinear warp resultant of previous nonlinear registration of the structural scan to the MNI T1 template (5) regression of motion parameters, mean white matter and cerebrospinal fluid signals; (6) smoothing with a 8-mm full-width half-maximum Gaussian kernel to decrease spatial noise and (7) high pass temporal filtered (filter width of 128s).

Final images were visually inspected to guarantee that they had not any disproportionate head motion and that participants had no brain lesions. One participant was excluded due to head motion higher than 2 mm in translation and 1.5° in rotation. In addition, two participants were excluded due anatomical abnormalities and two participants were excluded due to difficulties understanding the content of the videos. The fMRI paradigm analyses were performed using the general linear model (GLM) approach.

Statistical analysis

The design of the study is based on two main effects: the effect of condition (self or other) and the effect of the vignette emotional valence (positive, negative and neutral). First, to address the different neural responses to positive and negative videos in both self and other conditions, four different directional t-contrasts were computed. These contrasts included: self positive > self neutral, self negative > self neutral, other positive > other neutral and other negative > neutral. Next, to assess each activation profile (e.g. activation in self condition when watching positive stimuli) one sample t-tests were computed. A family wise error (FWE) correction at voxel-level was adopted to correct for multiple corrections, with results being significant corrected for FWE<0.05.

In addition, the above-mentioned contrast images were entered in repeated measures ANOVA with two within-subjects factors: condition (self and other) and valence (positive and negative). T-contrasts and an F-contrast were computed to assess factors main effects and the interaction effect. Results of the main effect of valence were corrected for multiple comparisons using FWE at voxel-level and results were considered significant for false FWE<0.05. In addition, the results of the main effect of condition and interaction effect were thresholded at p<0.005 and cluster size >15 voxels. Small volume correction (SVC) with a threshold of P ≤ 0.05 was used to correct for multiple comparison. Regarding the SVC, the ROIs were selected from Bzdok et al., (2012), a meta-analysis that assesses both affective empathy and theory of mind regions based on 247 fMRI experiments, including a total of 1790 participants. Finally, a combination of visual inspection and Anatomical Automatic Labeling atlas (AAL) was used for anatomical labelling (Tzourio-Mazoyer et al., 2002).

Results

Self Condition

First, we looked at the brain's response to positive (versus neutral) and negative (versus neutral*)* stimuli, separately. In the self condition, the response to positive vignettes included brain activations in the bilateral middle and superior temporal gyrus (STG) – including the Helsch's gyri, TPJ, rolandic operculum, angular gyrus, postcentral, superior temporal poles (STP), and left posterior insula; bilateral precentral gyrus; bilateral SMA; left PCC and precuneus; bilateral ventral anterior thalamus; left medial dorsal thalamus; left IFG and left frontal inferior orbital extending to AI. Still in the self condition, the brain's response to negative vignettes included the bilateral middle and STG - including the Helsch's gyri, left TPJ, rolandic operculum, STP and left posterior insula; bilateral precentral gyrus; left SMA; left PCC & precuneus; bilateral inferior occipital gyrus; left fusiform, left ventral anterior thalamus and left IFG. Results are displayed in Figure III – 2 and Table III – 1.

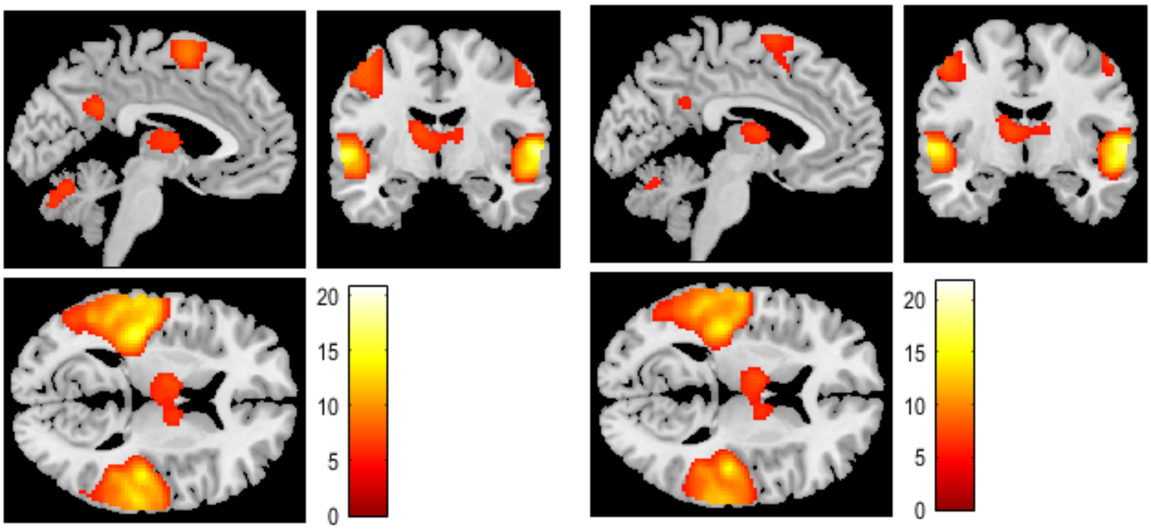


Figure III- 2 Activations identified in the self-condition when hearing to (left) positive (versus neutral) contents or (right) negative (versus neutral) contents expressed by their partner.

Table III- 1. Regional brain activity in the self condition for positive (> neutral) and negative (> neutral) stimuli.

		MNI coordinates				
	Region of interest	x	y	z	Z	Cluster
Positive > neutral						
L	Superior temporal gyrus	-58	-20	2	>8	5509
R		56	-24	-2		4391
L	Precentral gyrus	-40	-4	48	7.22	728
R		52	-2	52	6.65	288
L	Supplementary motor area	-2	4	66	6.57	512
R		4	12	64		
L	Posterior cingulate & precuneus	-10	-50	30	6.14	301
L	Ventral anterior thalamus	12	-6	14	5.66	631
L	Medial dorsal thalamus	-2	-12	8	5.19	
R	Ventral anterior thalamus	10	-4	12	5.18	
L	Frontal inferior orbital / Insula	-36	26	-2	4.78	76
L	Inferior frontal gyrus	-50	18	18	4.77	50
Negative > neutral						
L	Superior temporal gyrus	-58	-20	2	>8	4764
R		56	-24	-2		4090
L	Precentral gyrus	-40	-4	48	7.09	725
R		52	0	52	6.91	230
R	Inferior occipital gyrus	44	-80	-8	6.05	80
L		-50	-76	-2	5.62	
L	Fusiform	-42	-76	-18	5.82	254
L	Ventral anterior thalamus	10	-6	12	5.89	471
L	Supplementary motor area	-2	4	66	5.81	389
L	Posterior Cingulate & precuneus	-10	-50	30	5.17	73
L	Inferior frontal gyrus	-48	18	20	4.66	39

Other Condition

Regarding the other condition, the neural response to positive vignettes (versus neutral) included the bilateral middle and STG - including the helsch’s gyri, TPJ, angular gyrus, left SMG, posterior insula, STP and left rolandic operculum; left precentral gyrus; left PCC and precuneus; left IFG, bilateral middle/ inferior occipital gyrus; left temporal inferior; bilateral ventral anterior thalamus; left medial dorsal thalamus; left midbrain and left frontal inferior orbital extending to AI.

Furthermore, the brain’s response to negative vignettes (versus neutral) included activations in the bilateral middle and STG - including the helsch’s gyri, TPJ, angular gyrus, middle and right STP, postcentral, SMG and left rolandic operculum; left precentral gyrus; left PCC & precuneus, bilateral OI, left fusiform; left inferior temporal gyrus; left frontal inferior operulum and left frontal inferior orbital extending to anterior insula. Results are displayed in Figure III - 3 and Table III - 2.

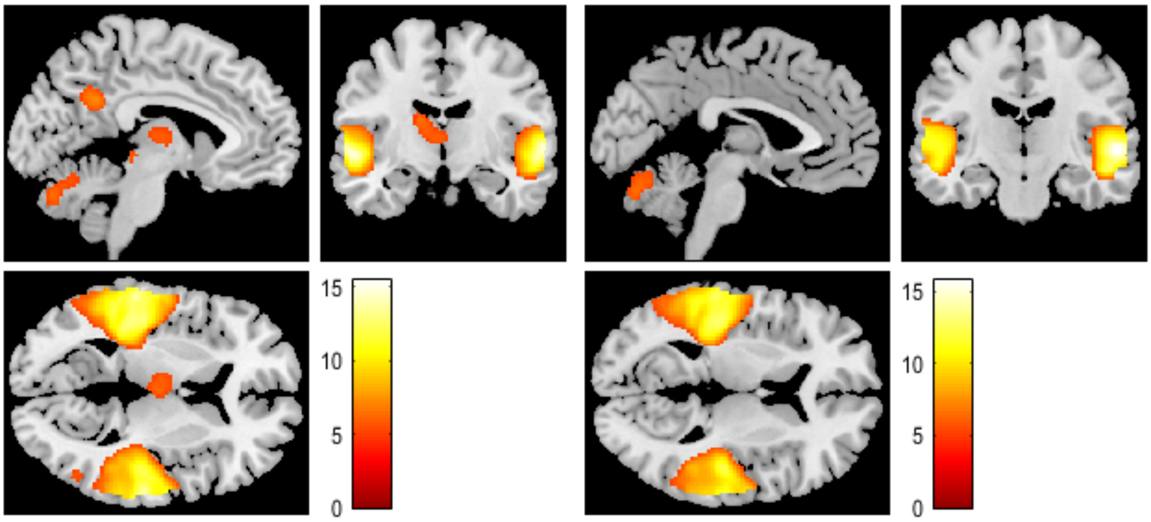


Figure III - 3 Activations identified in the other-condition when hearing to a) positive (versus neutral) contents or b) negative (versus neutral) contents expressed by their partner.

Table III - 2. Regional brain activity in the self condition for positive (> neutral) and negative (> neutral) stimuli.

		MNI coordinates				
	Region of interest	x	y	z	Z	Cluster
Positive> Neutral						
L	Superior temporal gyrus	-56	-8	-4	>8	4951
R		58	-14	0	>8	3788
L	Precentral gyrus	-40	-4	46	6.01	116
L	Posterior cingulate & precuneus	-10	-52	30	5.71	231
L	Inferior frontal gyrus	-48	14	22	5.50	109
R	Inferior occipital gyrus	44	-80	-8	5.45	15
L	Inferior temporal gyrus	-46	-50	-18	5.30	325
L	Inferior occipital gyrus	-44	-77	-10	5.12	
L	Ventral anterior thalamus	-14	-10	16	5.18	270
L	Medial dorsal thalamus	-2	-12	8	4.99	
L	Frontal inferior orbital /Insula	-36	24	-4	5.14	100
L	Midbrain	-4	-28	-4	5.00	16
R	Ventral anterior thalamus	12	-8	14	4.71	26
Negative > neutral						
L	Superior temporal gyrus	-56	-8	-4	>8	4440
R		58	-14	0	>8	3618
L	Inferior occipital gyrus	-44	-77	-10	5.35	225
L	Fusiform	-42	-71	-18	5.12	
L	Precentral gyrus	-40	-4	46	5.15	37
L	Inferior temporal gyrus	-42	-46	-14	5.11	32
L	Posterior cingulate & precuneus	-10	-52	30	4.99	67
L	Inferior frontal gyrus	-46	14	22	4.91	42
L	Frontal inferior orbital / Insula	-40	-18	-15	4.72	18

Comparing Self/Other Condition

Concerning the results of the ANOVA, the first factor (condition) showed significant differences in brain activity in both self > other and other > self directions. The contrast self > other evidenced significant clusters in the activity of the bilateral STC and insula. On the other hand, the contrast other>self revealed several clusters, namely in the bilateral fusiform, right middle occipital gyrus, left supramarginal/ postcentral (including insula), left inferior temporal gyrus, left precentral gyrus, left angular gyrus, and right temporal middle/ TPJ. The summary of the results is displayed in Table III - 3 and Figure III – 4.

Table III -3. Differences in brain activation between self and other conditions; Voxel Threshold: p <= .005; Cluster Size Threshold: k >15.

		MNI coordinates				
	Region of interest	x	Y	Z	Z	Cluster
	Self > Other					
L	Superior temporal gyrus	-54	-10	0	4.21	355
L	Insula	-44	-14	2	4.08	
R	Superior temporal gyrus/Insula	48	-8	4	3.69	82
	Other > Self					
L	Fusiform	-26	-34	-18	3.82	61
R	Middle occipital gyrus	38	-74	8	3.77	134
R	Fusiform	28	-36	-16	3.55	101
L	Supramarginal gyrus /Postcentral	-54	-20	18	3.23	142
L	Inferior temporal gyrus	-44	-50	-6	3.19	102
L	Precentral gyrus	-46	4	18	3.14	86
L	Angular gyrus*	-36	-62	30	3.12	53
R	Temporal Middle*	50	-48	18	2.87	32
	Interaction effect					
L	Caudate	-18	-18	24	3.71	425

*Regions that survived SVC at p > 0.05 FWE with a 10 mm sphere using coordinates from (Tzourio-Mazoyer et al., 2002)

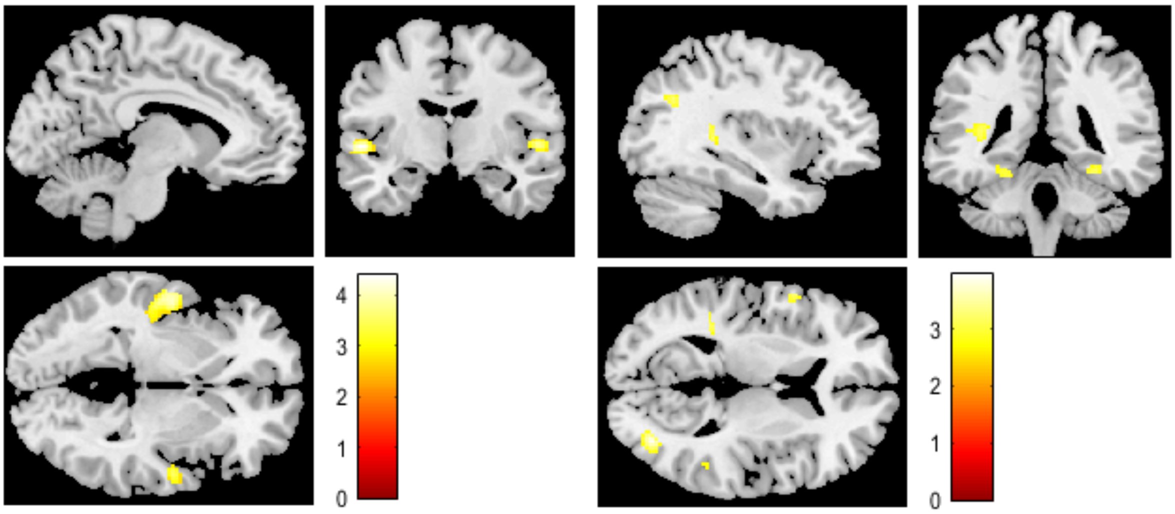


Figure III- 4 Activations identified when hearing to contents expressed by their partner for a) self-condition higher than the other condition or b) other-condition higher than the self-condition.

Regarding the second factor (valence), differences were only significant in the contrast positive > negative video vignettes. Indeed, the brain's response to positive vignettes (versus negative) included brain activations in the left postcentral gyrus (including IPL) and right SMA. The summary of the results is displayed in Table III - 4.

Table III - 4. Differences in brain activation for positive (> negative) contents expressed by their partner; Voxel Threshold: p <= .001; FWE corrected; Cluster Size Threshold: k >15.

		MNI coordinates				
	Region of interest	x	y	z	Z	Cluster
L	Postcentral gyrus	-34	-24	50	6.10	182
	Inferior parietal lobe	-46	-24	38	5.24	
R	SMA	6	-14	48	4.82	45

Finally, an interaction effect between condition (self and other) and valence (positive and negative) was found in the left caudate with increased functional activity in the self condition for negative vignettes (see Table III - 3).

Discussion

In this study, we looked at the neural basis of self /other processing in the context of couple interaction during an fMRI task. In this task, each participant watched a set of video-vignettes of a previously recorded real interaction with his/her partner and was asked to focus either on what he/she was feeling (self condition) or on what his/her partner was feeling during the video (other condition). Consistent with our hypothesis, our results revealed a significant, although not complete, brain activity overlap between self and other conditions.

Both self and other conditions recruited brain regions associated with socio-affective processing (e.g., AI, anterior thalamus, MTG, IPL and IFG and PCC), as well as with socio-cognitive processing (e.g., TPJ, precuneus and the STP (Bzdok et al., 2012; Schurz et al., 2014). The fact that both pathways of social processing were recruited, provides support to the idea of an interaction between affective and cognitive systems when understanding both ours and another's internal states (Lamm et al., 2010; Zaki et al., 2009).

Furthermore, the almost complete overlap between self and other neural activations brings support to the idea that understanding our own internal states and those of others requires similar psychological processes and associated brain mechanisms. This is consistent with the ‘‘simulation’’ theories of social cognition which argue that to understand the internal states of a close other we “simulate” them in ourselves (Adolphs, 2002; Vittorio Gallese & Goldman, 1998), engaging our own brain and body functions (Lamm et al., 2016; Rutgen, Seidel, Rie ansky, & Lamm, 2015). Indeed, the closer the other is to oneself, the more likely we are to base inferences about them on knowledge about oneself (Adolphs, 2002; Vittorio Gallese & Goldman, 1998).

Regarding the differences between conditions, while the self-condition was associated with higher activations in the bilateral STG region including the posterior insula, the other-condition was associated with higher activations in the bilateral fusiform, right middle occipital, left SMG, left temporal inferior, left precentral, left angular gyrus, and right temporal middle/ TPJ. As expected, the self-condition engaged brain regions traditionally linked with interoceptive processing and affect sharing, such as the posterior insula (Chang et al., 2012; Craig, 2003; Singer et al., 2004). Previous work had reported increases in activation of the insula and the anterior cingulate cortex for self-related judgments when compared to other related judgments (e.g., Qin & Northoff, 2011).

On the other hand, the other-condition recruited to a further extent brain areas involved in emotion recognition, such as the left inferior temporal gyrus, the bilateral fusiform and the left occipital middle (Adolphs, 2002; Hooker, Verosky, Germine, Knight, & D’Esposito, 2008; Kawasaki et al., 2012) as well as in the representation of another’s internal states/TOM, like the angular gyrus and the TPJ (Bzdok et al., 2012; Schurz et al., 2014). Importantly, the most significant region (the only one that survived to multiple comparison correction) for this contrast (other >self) was the right TPJ, a brain area thought to have a major role in differentiating self and other perspectives (Santiesteban, Banissy, Catmur, & Bird, 2012; Steinbeis, 2015).

Indeed, the ability to distinguish and switch between self and other perspectives plays a key role in social cognitive processes (Santiesteban, White, et al., 2012). Thus, the higher activation of the right TPJ in the other condition suggests that, in order to assume their partner’s perspective, participants had to inhibit their own perspective (Steinbeis, 2015).

Furthermore, an interaction effect between emotional valence and condition was found in the caudate nucleus. Specifically, we found a higher activation in the left ventral caudate in the self-condition when participants were watching negative vignettes. Although the role of the caudate nucleus for self-referential processing remains unclear, previous evidence has shown that the left ventral caudate region is functionally connected to the posterior insula (Di Martino et al., 2008; Huang et al., 2017), a region known for interoceptive processing and affect sharing (Craig, 2003; Singer et al., 2004). In line with this result, recent evidence has shown that the functional connectivity between the posterior insula and the ventral caudate was significantly correlated with affective functions (Huang et al., 2017).

Finally, regarding the valence of the video-vignettes, in our study the neural response to positive vignettes was evident than the responses to the negative ones. This increased activation was found for several brain areas, namely in the right SMA and left postcentral gyrus including the IPL. These findings are not in line with previous data suggesting that humans have a remarkable ability to share the distress of others, but may react less to their joy (Perry, Hendler, & Shamay-Tsoory, 2012). What our results seem to show is that when interacting with a close other such as an intimate partner, people may react more to positive emotions expressed of the other. In addition, our results showed that different brain areas were recruited for positive and negative stimuli. Specifically, positive contents were associated with higher activation in the bilateral thalamus, while negative stimuli were associated with higher activation in the inferior occipital gyrus and the fusiform gyrus – associated with self-referential processing of sad stimuli (Reniers, Völlm, Elliott, & Corcoran, 2014). These results are in line with previous findings showing that depending on the valence of the stimuli, different brain areas are recruited (Phan, Wager, Taylor, & Liberzon, 2002).

In summary, the present results show that there is an almost complete overlap between the neural systems that underlie processing our own internal states and those of a close other. In addition, the other condition recruited additional brain regions involved in the representation of another's internal states, TOM and self-other distinction. These findings provide support to simulationist accounts of social cognition which posits that our capacity to understand another's mind relies on our privileged access to our own mental states (Lamm et al., 2016; Preston & de Waal, 2002; Rutgen et al., 2015). This may clarify why individuals with higher self-awareness are also more capable of understanding the other´s emotions. This was recently shown by Böckler, Herrmann, Trautwein, Holmes, and Singer, (2017) that investigated the role of a training-induced understanding of oneself for the enhanced understanding of others. Results revealed that the degree to which participants improved their understanding of themselves predicted their improvements in high level ToM performance.

The results of this study have clear implications not only for couples functioning, but also for other human dyads such as mother-infant or helping professional relationships like physician-patient. They are also important for the understanding of clinical disorders in which the mechanisms of self and other processing are altered, such as depression or autism.

Future studies may build on the present results and move further to use analytic methods such as DCM that explicitly measure the influence that one neural system exerts over another. DCM analysis allows us to understand the flow of information within emotional and cognitive networks, as well as between networks. For example, the knowledge of the information flow between socio-affective and socio-cognitive networks, will clarify if these networks are hierarchically related, with the ability to abstract mental state attributions being dependent on the ability to simulate the other state.

Author Contributions

J.Coutinho, J.Decety and O.F. Gonçalves designed the study concept and design. J.Coutinho and P.O-Silva collected data for the experiments. S.Esmenio performed the data analysis and interpretation under the supervisison of J.Coutinho and J.M.Soares. S.Esmenio, J.Coutinho and J.M.Soares wrote the paper and all authors reviewed and approved the final draft.

Acknowledgements

This study was funded by BIAL Foundation (Grant number 87/12); by the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Education and Science through national funds and co-financed by FEDER through COMPETE2020 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653); by the Portuguese Foundation for Science Doctoral scholarship (PD/BD/105963/2014).

CHAPTER IV[3]

“Changes in the effective connectivity of the social brain when making inferences about self and close others”

---

[3] This Chapter reproduces the publication:

Esmenio, S., Soares, J.M., Oliveira-Silva, A., Gonçalves, Ó.F., Friston, K., Coutinho, J., (2019), Understanding other´s internal states in the context of romantic relationships: an effective connectivity study, Human behavior (submitted).

Abstract

Studies of the neural basis of social cognition have shown that the ability to infer our own and other's mental states relies on common brain pathways; particularly in the case of close relationships (e.g., romantic relationships). Despite the evidence for shared neural representations of self and others, less is known the distributed processing within these common neural networks. In particular it is not known whether there are specific patterns of internode communication when focusing on other versus the self. The aim of the present study was to characterize context-sensitive coupling among social brain regions involved in self and other understanding. Forty-two participants – in a committed monogamous relationship – underwent an fMRI while watching emotional video vignettes of their partner. Participants were asked to elaborate on their partner´s experience (other-condition) or on their own experience (self-condition). We used dynamic causal modelling (DCM) to quantify the associated changes in effective connectivity (EC) in a network of brain regions with an established role in social cognition. This network included the bilateral temporoparietal junction (TPJ), the posterior cingulate (PCC)/precuneus (prec) and the bilateral middle temporal gyrus (MTG). DCM revealed that: (1) the PCC plays a central/coordination role, (2) the bilateral MTG receives driving input, suggesting that stimulus-bound information is first processed in a language (and narrative comprehension) region; (3) the right TPJ evinced a selective increase in its sensitivity when focusing on the other's experience, relative to focusing on oneself.

Keywords: Social cognition; Self and other; DCM; PEB; Romantic relationships

Introduction

Social neuroscience research has shown that when trying to understand another's emotional and mental states, we rely on psychological processes and brain systems similar to those that we use to understand our own internal states (e.g. Lamm et al., 2016). Similar brain networks are recruited when processing self and other´s internal states in both affective (e.g. Jackson et al., 2005; Lamm et al., 2011; Singer et al., 2004), and cognitive tasks (e.g. Lombardo et al., 2009; Mitchell et al., 2006; Ochsner et al., 2004) . Consistent with these findings, – and supporting the idea that inferring or understanding oneself and others are only “semi-independent skills" (Dimaggio et al., 2008) – recent evidence suggests that enhancing the capacity to understand our own thoughts and feelings increases the ability to infer those of others (Böckler et al., 2017).

According to simulation theories of social cognition, the closer the other is to oneself, the more likely we are to ground inferences about them on knowledge about oneself (Adolphs, 2002; Aron et al., 1991; Vittorio Gallese & Goldman, 1998). Therefore, close relationships (such as romantic partners) suggest themselves as a relevant context to characterize the intimate relationship between self and other processing. A previous study, in which participants were presented video-vignettes of their romantic partners, confirmed a significant overlap between the functional anatomy of self and other processing, implicating brain regions associated with both socio-affective and socio-cognitive systems (Esménio et al., 2019).

Despite the evidence for shared neural representations of self and other (e.g. Lombardo et al., 2009), less is known about how information flows within these common neural networks, specifically, are there specific patterns of internode connectivity when focusing on other versus on the self? The objective of the present study was to characterize the information flow among social brain regions involved in self and other understanding. We therefore analysed the effective connectivity (EC) in a network of brain regions implicated in social cognition that included the bilateral temporoparietal junction (TPJ), the posterior cingulate (PCC)/precuneus (prec) and the bilateral middle temporal gyrus (MTG). A previous study has established that these brain regions are engaged by the experimental paradigm used in the present study (Esménio et al., 2019).

The role of each of these regions for social processing is well documented. For example, the MTG is primarily involved in language related processes, such as semantic processing and speech perception (Amft et al., 2014; Spreng, Mar, & Kim, 2009), story and narrative comprehension (Mar, 2011), memory processing; particularly, autobiographical memory (Spreng et al., 2009) and emotional processing (Leonhard Schilbach et al., 2012). The PCC/prec is involved in a wide range of highly integrated tasks (Cavanna & Trimble, 2006), being associated with both self-related processes including self-representation and self-reflection (Cavanna & Trimble, 2006; Johnson et al., 2006), consciousness (Northoff et al., 2006; Vogt & Laureys, 2005), future thinking and prospective memory (Christoff & Gordon, 2008), and other-related processes, such as mentalizing/TOM (Bzdok et al., 2012; Rebecca Saxe & Powell, 2006; Schurz et al., 2014); narrative comprehension (Mar, 2011) as well as in empathic and forgivability judgments (Farrow et al., 2001; Ochsner et al., 2004).

Finally, the TPJ is – a key social brain region – is traditionally associated with social-cognitive processes such visual perspective taking, and mental inference /TOM (Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006; Kanske et al., 2015; Ramsey, Hansen, Apperly, & Samson, 2013; Rebecca Saxe & Kanwisher, 2003; Schurz et al., 2014; Van Overwalle, 2009). Importantly, whereas the left TPJ (LTPJ) seems to be mainly involved in language processes (Binder, Desai, Graves, & Conant, 2009) and intention detection (Berthoz, Armony, Blair, & Dolan, 2002), the right TPJ (RTPJ) has been mostly associated with self-awareness (Jean Decety & Lamm, 2007), and differentiation between self and other perspectives (Santiesteban, Banissy, et al., 2012; Steinbeis, 2015). Indeed, the RTPJ has been shown to play a major role during ToM (Rebecca Saxe & Kanwisher, 2003), especially when a difference in perspective exists between self and other (Aichhorn et al., 2006; Santiesteban, Banissy, et al., 2012; Sommer et al., 2007). In short, the right TPJ appears to play a crucial and context-sensitive role in functional integration, when making inferences about others, relative to self.

Despite the converging findings supporting the involvement of these brain areas in social understanding, the functional integration among these regions has not been established. In particular, the changes in coupling between (i.e. extrinsic connectivity) and within (i.e., intrinsic connectivity) that underwrite differential processing during inferences about self and other are largely unknown. Analytic approaches such as dynamic causal modelling (DCM) are important tools to characterize the causal relationships between the brain nodes of a network.

Dynamic causal modelling (DCM) is a generative model-based Bayesian approach that infers the EC among distinct brain regions using forward models of how neuronal states influence each other – and detailed biophysical hemodynamic models of how neuronal states generate fMRI signals (Friston, 2009; Friston, Harrison, & Penny, 2003). More specifically, DCM supplements a forward model of how cortical regions interact with a hemodynamic model that transforms neuronal activity into the measured response (e.g., blood oxygen level response) (Friston, Kahan, Biswal, & Razi, 2014; Razi, Kahan, Rees, & Friston, 2015). This combination of an a priori biologically plausible neural network model with the measured BOLD response, makes it possible to infer information flow in terms of directed intrinsic and extrinsic effective connectivity; namely the effect that one neuronal system has on another (Daunizeau, David, & Stephan, 2011; Friston, 2009).

As the method of choice for modelling causal interactions in neuroimaging data, task-related DCM has been extensively applied to study a wide variety of processes in both clinical and non-clinical populations. Some examples include DCM studies on speech perception (Osnes et al., 2011), motor functioning (Minkova et al., 2015), language and motor rehabilitation post-stroke (Kiran et al., 2015; Rehme, Eickhoff, Wang, Fink, & Grefkes, 2011), attention (Fairhall, Indovina, Driver, & MacAluso, 2009), executive function in major depression (Schlösser et al., 2008), spatial and lexical processing (Deng, Guo, Ding, & Peng, 2012), response inhibition and working memory in Schizophrenia (Allen et al., 2010; Zhang et al., 2013), inter-hemispheric integration in Alzheimer's disease (Rytsar, Fornari, Frackowiak, Ghika, & Knyazeva, 2011).

In this study, we used DCM combined with Parametric Empirical Bayes (PEB) and Bayesian model reduction (BMR) to examine the changes in EC associated with focusing on the partner’s internal states, in comparison to focusing on one's own experiences. Technically, this entails inverting a fully connected model for each subject, using subject-specific posteriors over the model parameters to estimate group means (using PEB) and then removing redundant parameters (using BMR). In our application, the key parameters of interest were changes in intrinsic and extrinsic connectivity that model the context sensitive changes in coupling or information flow due to focus on other, versus self. Following previous work using this experimental paradigm (Esménio et al., 2019), our primary hypothesis was that focusing on the partner (other condition) would increase in the sensitivity of RTPJ to its afferents from other nodes within the social brain network under study.

Method

Participants

Forty-two participants – in a monogamous romantic relationship for at least one year – were enrolled in this study. Prior to any procedure, all participants were screened on the telephone for assessment in relation to inclusion/exclusion criteria. Inclusion criteria included: (1) the absence of any diagnosed neuropsychiatric or neurodegenerative disorder; (2) absence, in the past year, of a dependency/abuse of alcohol or drugs; (3) ability to attend the MRI session; (4) age between 20 and 50 years. In the final sample, all the participants were Caucasian, right-handed and the age of the participants ranged from 23 to 40 years old (M = 31.17, SD = 4.748; for men: M = 32.13, SD = 4.893, for women: M = 30.22, SD = 4.502). The duration of the relationships ranged from 1 to 15 years (M= 7.78 years, SD=4,76).

The study complied with the principles expressed in the Declaration of Helsinki (with the amendment of Tokyo 1975, Venice1983, Hong Kong 1989, Somerset West1996, Edinburgh 2000) and was approved by the University Institutional Review Board. At the beginning of this study, the procedure was explained to the participants who provided informed written consent.

Experimental Task

Each participant watched a set of videos of his/her romantic partner expressing personal contents, and was asked to, while watching, either focus on his/her own experience (self condition) or on his/her partner's experience (other condition). Twenty seconds video-vignettes were extracted from a previously video-recorded interaction task performed in a lab, where participants shared things that they either like (positive videos) or dislike (negative videos) about their partner (more details regarding the interaction task can be found in (Coutinho et al., 2017; Coutinho et al., 2018).

The final fMRI task comprised two blocks (self and other), each containing three different conditions: positive trials (N=8), negative trials (N=8) and neutral trials (N=6). The latter were extracted from the Emotional Movie Database (Carvalho, Leite, Galdo-Álvarez, & Gonçalves, 2012). Each trial consisted of: (1) fixation cross (5s); (2) instructions in accordance with the current block (3 s); (3) video vignette (20s); (4) behavioral response (4s). An example of a trial is shown in Figure 1. Blocks were displayed in a randomized order across participants and stimuli were displayed in a pseudo-randomized order across blocks.

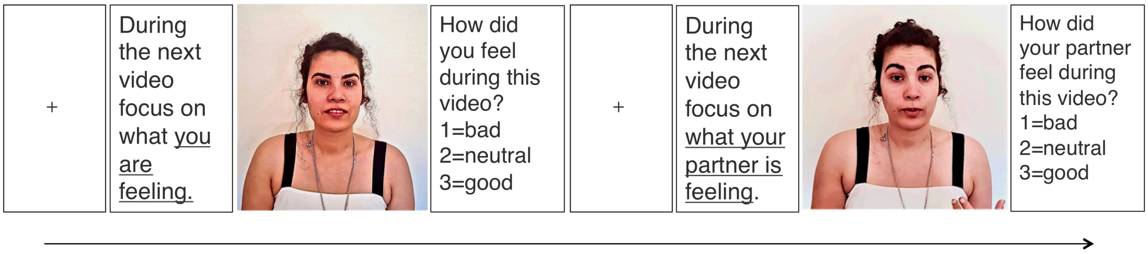


Figure IV -1 Scheme of a self and other emotional trial.

Image acquisition

Structural (T1) and functional images (T2*) were acquired with a clinically approved 3T MRI scanner (Siemens Magnetom Tim Trio, Erlangen, German). Data were acquired from each participant in a session that included one structural T1 scan (192 sagittal slices, repetition time (TR) =2000 ms.; echo time (TE) = 2.33 s, flip angle = 7°, slice thickness = 0.8 mm, slice gap = 0 mm, pixel size = 0.8 × 0.8 mm^2, field of view (FoV) = 256 mm) and one functional BOLD sensitive echo-planar imaging (EPI) sequence (39 axial slices; repetition time (TR) =2000 ms.; echo time (TE) = 29 ms., flip angle (FA) = 90◦, matrix size = 64 × 64, slice thickness = 3 mm, pixel size = 3 × 3 mm^2, field of view (FoV) = 222x222 mm). Synchronisation between the paradigm and the acquisition was insured for each TR.

Data preprocessing

Data preprocessing was performed using the Statistical Parametric Mapping software (SPM12; Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk). Preprocessing steps included: (1) slice-timing correction, (2) motion correction through the re-alignment to the mean image; (3) rigid-body registration of the mean functional image to the T1; (4) normalization of the functional acquisition to the Montreal Neurological Institute (MNI) standard space (Ashburner & Friston, 1999) through the application of a rigid body transformation and a nonlinear spatial normalisation following nonlinear registration of the T1 to the MNI T1 template (5) regression of motion parameters, white matter (WM) and cerebrospinal fluid (CSF) signals; (6) smoothing with a 8-mm full-width half-maximum Gaussian kernel to decrease spatial noise and (7) high pass temporal filtering (filter width of 128s) to remove low frequency noise. A general linear model (GLM) was inverted for each subject to identify subject specific regional responses for subsequent DCM analysis. All images were inspected visually to ensure that participants had no brain lesions or disproportionate head motion. Nine participants were excluded: 1 due to head motion higher than 2 mm in translation and 1.5° in rotation; 2 due to anatomical abnormalities; 2 due to technical problems; 4 due to abnormal patterns of activation during the video condition.

Effective connectivity analysis

In brief, our fMRI experimental design was a standard block design with a 2 x 3 factorial structure (a two level *self* versus *other* factor, and three levels of valence; *positive*, *negative* and *neutral*). For the DCM analyses, we focused on the main effects of the self-other factor. In other words, we asked whether this factor changed directed connectivity within the social network under study.

DCM was implemented using the Statistical Parametric Mapping software (SPM12; Wellcome Department of Cognitive Neurology, London, UK) to estimate the effective connectivity between the regions of interest (ROIs). Based on the group results from a previous study (Esménio et al., 2019), we selected the brain regions associated with high-level social processing (Alcalá-López et al., 2017; Bzdok et al., 2012; Schurz et al., 2014); namely, those regions activated during both self and other condition. These included the left and right MTG, the PCC/precuneus and the left and right TPJ. Since the RTPJ was the only significant region found in the previous study for the contrast self versus other this region was designated as the index node, who was connectivity dependent upon the self-other context.

Regarding the regional responses for each subject (i.e. the selection of the ROIs), five peak coordinates were used: LMTG (-56, -14, -12), RMTG (56, -10, -12), PCC/prec (-10, -52, 30), LTPJ (-49, -61 28) and RTPJ (50, -48, 18). These coordinates were obtained by combining several group analysis from a previous study (Esménio et al., 2019), and standard coordinates from the recent literature on social brain (Alcalá-López et al., 2017). The regional responses corresponded to the principal eigenvariate within an 8-mm sphere, centred on the corresponding subject-specific peak activation within 10 mm of the group coordinates. Three subjects were excluded because no significant activation was observed in the PCC.

Next, a DCM model was computed for each subject, combining: (1) the five regions of interest; (2) a fully connected model (displayed in Figure 2 a), where every region shares a connection with all the other regions in the network (i.e., full extrinsic connectivity) and has a connection to itself (i.e., intrinsic connections), (3) driving inputs on every node (apart from the index node – the RTPJ), where the driving input comprised the visual stimulation during the vignettes viewing blocks (displayed in Figure 2 b); (4) condition-specific or modulatory effects on all our friends to the RTPJ (including intrinsic self-connections, as shown in Figure 2 c). In summary, this 5 ROI model comprised a fully connected architecture (i.e. 5 intrinsic/self-connections and 20 between regions/extrinsic connections), 4 driving inputs exerting direct effects on four ROIs - the LMTG, the RMTG, the PCC and the LTPJ - and 5 context-sensitive or modulatory effects (i.e., the four other ROIs that shared a connection with the RTPJ – LMTG, LTPJ, PCC and RMTG - and a self-modulatory effect). The ensuing model architecture is displayed in Figure 2 d.

The selections of the position for the modulatory effect and the driving inputs' locations was based on the results of the previous study and the connectivity architecture in (Alcalá-López et al., 2017). The latest represents a meta-analysis that derives a social brain definition from 26 meta-analyses of social-cognitive capacities with significant convergence from original 25,339 initial foci from 3972 neuroimaging studies in 22,712 participants.

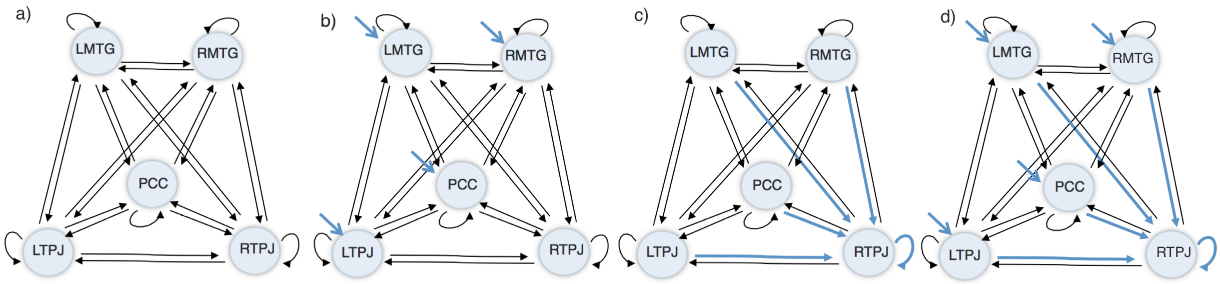


Figure IV - 2 – DCM initial model. a) Connectivity architecture. b) Driving inputs. c) Modulatory effects. D) Final model.

After the model specification, this model was estimated and inverted for each subject, and the ensuing posterior densities over connectivity parameters (i.e. posterior means and covariances) were taken to the between subject level for inference about group effects using

Parametric Empirical Bayes (PEB). PEB random effects (RFX) assumes that all subjects use the same model architecture but express different parametric effects in terms of the connection strengths or their modulation. In other words, all subjects share the same architecture but express condition-specific effects to a greater or lesser extent. The final step, the elimination of redundant parameters using Bayesian model reduction (BMR), enables one to identify context-sensitive changes in connectivity by comparing models that do and do not contain modulatory parameters.

In fact, three different PEBs (and consequent BMRs) were performed, each addressing one sort of connectivity: i.e. the average connectivity across conditions (corresponding to DCM's matrix A); the driving inputs (matrix C – the blue arrows in Figure IV - 2 b); and the context-sensitive or modulatory effect (matrix B – the blue arrows in Figure IV - 2 c). The results of this reductive form of Bayesian model selection are shown in Figure 3 and 4, where the parameter estimates of the extrinsic connections correspond to the underlying connection strengths or information flow (in Hertz). In respect to intrinsic connections, the effects are modelled in terms of the log scaling of inhibitory self-connections (of .5 Hz). Only connectivity parameters that survived to a posterior probability of 95% are shown (when comparing models with and without each parameter).

Results

Regarding the connectivity architecture, a quantitative distinction must be made between the extrinsic and intrinsic or self-inhibition connections: extrinsic connections represent the effects (in Hz) that one region has on another, while intrinsic connections represent a log scaling of self inhibition. Self-inhibition represents the rate of decay of neuronal activity, where a higher self-inhibition means a region is less sensitive to its inputs. Similarly, a decrease in self-inhibition or intrinsic connectivity corresponds to a disinhibition or increased in sensitivity to afferent input. Note that this intrinsic disinhibition increases sensitivity to all afferent inputs; whereas an increase in extrinsic afferent connectivity is specific to the extrinsic connection in question.

In terms of average connectivity a number of connections were removed: from LTPJ to PCC; from RMTG to LMTG and to PCC; and from RTPJ to LTPJ, to PCC and to RMTG. In the final connectivity architecture: (1) the PCC excerts a positive influence in every region but only receives input from LMTG, (2) the RTPJ receives input from all the other regions having only one efferent to LMTG, (3) the LTPJ exerts a negative influence on all the other nodes (except for PCC), which can be interpreted as a “tonic” inhibition (Alcalá-López et al., 2017), (4) in both bilateral regions; i.e. MTG and TPJ, the information flows from the left region to the right node. In respect of the intrinsic connections, we can see that there was a “tonic” negative self-inhibition in all the nodes in this network, particularly in the PCC (p =0,91). The resulting connectivity architecture is summarised in Figure IV - 3.

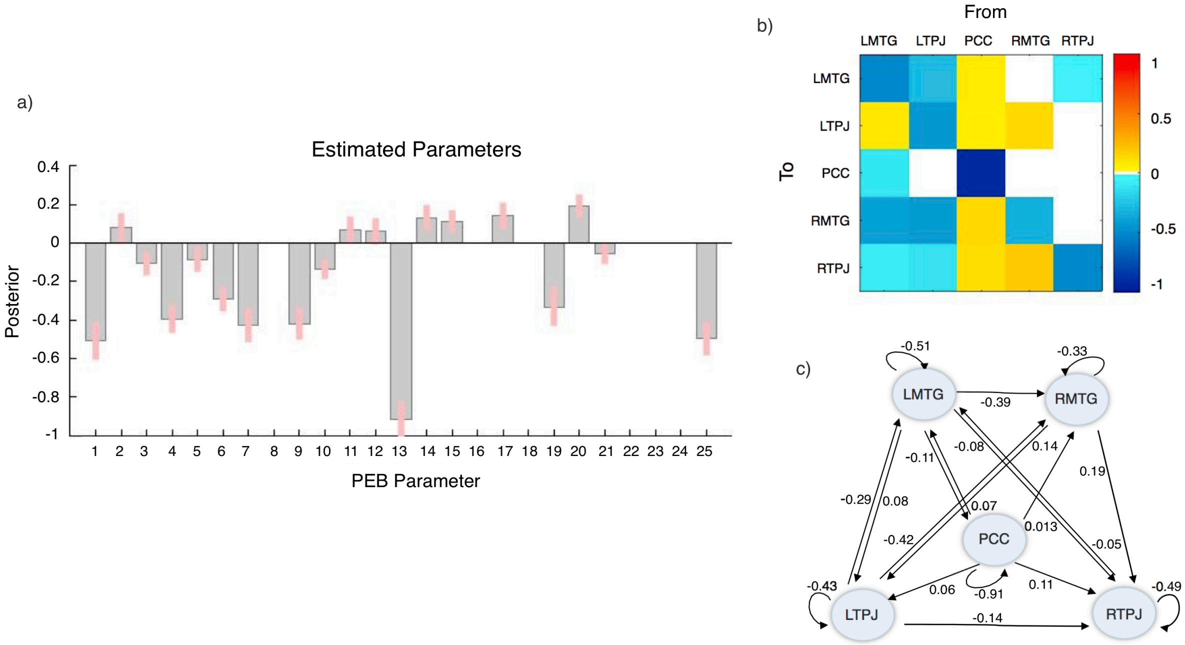


Figure IV - 3 - Average or “Baseline” Connectivity results. a) Parameters posterior estimates. b) Connectivity matrix. c) The structure and parameters of the winning model. The black lines/values illustrate the (natural) connectivity between brain regions; i.e., irrespective of stimulus and task. The numbers are the strength of connectivity (Hz).

In terms of driving inputs, the winning model retained only driving inputs to the bilateral MTG (left, Ep =0,086Hz and right, Ep=165Hz). Finally, in terms of the context sensitive changes in connectivity, the only modulatory effect that survived Bayesian model reduction was a decrease in RTPJ’s self-inhibition. (Ep=-0.91). In other words, RTPJ was disinhibited during the other condition; thereby increasing its sensitivity to all its afferents. These results are summarised in Figure IV - 4.

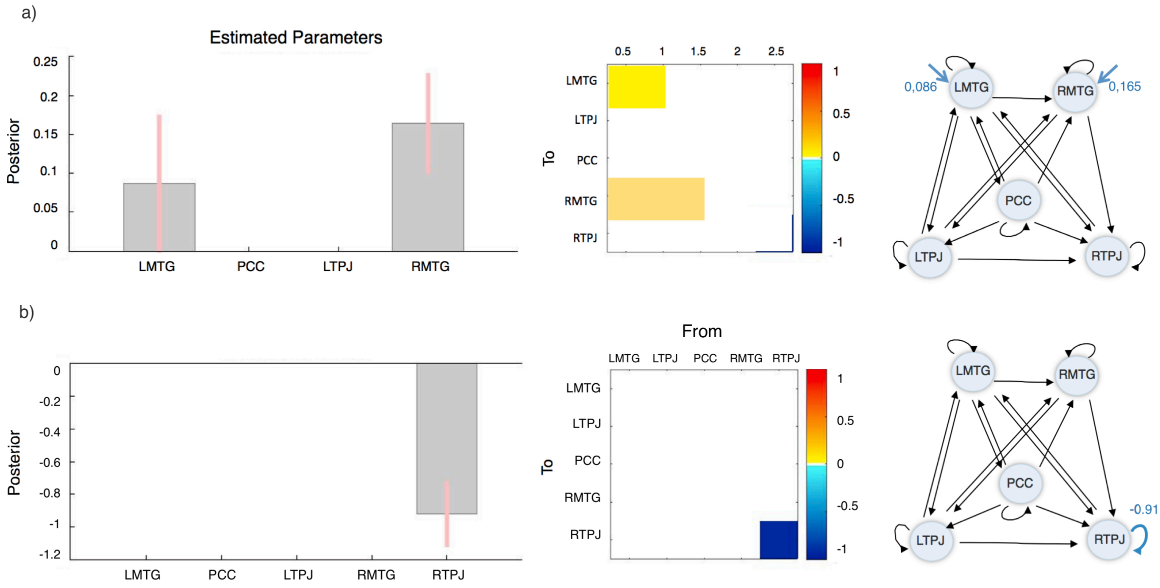


Figure IV - 4 – Driving inputs and modulatory effect results following Bayesian model reduction. a) Driving inputs. b) Modulatory effects or condition-specifics. (left) Parameters posterior estimates (EP). (center) Connectivity matrix. (right) The structure and parameters of the winning model. The black lines/values illustrate the connectivity between brain regions. The arrows in blue represent the driving inputs (upper) and modulatory effects (lower), respectively. The numbers quantify the strength of connectivity or information flow (Hz).

Finally, in order to take these results further, we performed a supplementary PEB analysis, focusing exclusively on the combination of intrinsic (within region) connections (i.e., all intrinsic connections were allowed to change). The results show that the disinhibition in RTPJ's (p=-0.64) sensitivity was accompanied by a decrease in LMTG's (Ep=0.22) intrinsic sensitivity. In other words, if all nodes are allowed to change their excitability, the differential activation elicited in RTPJ is explained by a reciprocal change in LMTG's and RTPJ's excitability. These results are shown in Appendix B in Supplementary Figure IV - 1.

Discussion

Previous research in social neuroscience of self and other processing speaks to the existence of shared neural systems for self and other processing (e.g. Lamm et al., 2016; Lamm et al., 2011; Lombardo et al., 2009; Singer et al., 2004). In particular, in the context of romantic relationships, a remarkable overlap was found in the brain regions recruited when attending to one's own internal states and those of a partner – as shown in our previous work (Esménio, Soares, Oliveira-Silva, Gonçalves, et al., 2019). This previous study also revealed that focusing on the partner preferentially recruited further brain regions involved in socio-cognitive processes, such as the RTPJ.

Hence, in the present study, we used DCM (combined with PEB and BMR) to estimate the information flow within a social brain network comprising the bilateral TPJ, PCC/precuneus and bilateral MTG, during a social inference task. We were especially interested in analysing changes in directed connectivity or information flow when participants focused on their romantic partner, rather than on themselves.

Our results showed that – in terms of extrinsic connections – in the final model a small number of extrinsic connections were redundant, in particular the connections to the PCC and from the RTPJ. Regarding the PCC, these results suggest that – as in the Default Mode Network which is also known as a mentalizing network – this region seems to have a coordination or orchestrating role within social brain networks (Deshpande, Santhanam, & Hu, 2011; Esménio, et al., 2019a; Hagmann et al., 2008; Raichle, 2015). In addition, based on findings of the analyses developed by (Alcalá-López et al., 2017), using fMRI task-constrained and task-unconstrained modalities to compute the "functional coupling" between 36 social brain seeds, the PCC is a plausible candidate for mediating the information flow between low level-limbic networks and high-level cognitive networks devoted to social processes.

On the other hand, the RTPJ appears to play the role of a receptor node within the network under analysis, as it shares afferents connections with all the nodes, having only an efferent connection with the LMTG. This result is particularly interesting when considering the role of this region in high-level cognitive processes, such as detection of intention, belief reasoning, perspective-taking and self-other distinction (Marcel Brass, Ruby, & Spengler, 2009; Ramsey et al., 2013; Santiesteban, White, et al., 2012).

Regarding the driving inputs – that correspond to the visual stimulation during the vignette blocks – we found that the only necessary driving inputs were those that entered through the bilateral MTG. Taking in consideration that the driving inputs were conveying visual information, this result goes in line with the study by (Alcalá-López et al., 2017), that suggests that the high-level and the visual-sensory social networks are connected through the bilateral MTG and the bilateral posterior superior sulcus (pSTS). Finally, this result suggests that the social information depicted in, or provided by, the stimuli may have first entered the system through language and narrative processing regions (Mar, 2011; Spreng et al., 2009) to be represented at the level, and then assimilated hierarchically by integrative/ high-level regions, such as PCC and TPJ.

At last, concerning the modulatory effect of focusing on other relative to self, even though we tested for models where the RTPJ could selectively increase its sensitivity to different afferents or inputs, we found that a sufficient explanation for our data was an increase in postsynaptic responsiveness – as mediated by an intrinsic disinhibition. It is generally thought that these changes in excitability rest upon fast synchronous interactions between inhibitory interneurons and pyramidal cells that express NMDA receptors (Mkael et al., 2018; R. J. Moran, Stephan, Dolan, & Friston, 2011; R. Moran, Pinotsis, Friston, Robinson, & Roberts, 2013). It is important to note however that in this DCM analysis, we do not identify the source of the neuromodulatory effects mediating the social process under study; we only identify the brain regions that constitute the targets of any context-sensitive modulation.

In summary, based on the well documented relationship between the selected regions, i.e. the PCC, TPJ and MTG, and social cognitive processing (Mar, 2011; Leonhard Schilbach et al., 2012; Schurz et al., 2014), our results characterise the different roles that each of these nodes may play within this social brain network. Regarding the PCC, similarly to the interpretation of resting state studies of DMN connectivity (Deshpande et al., 2011; Esménio et al., 2019a; Hagmann et al., 2008; Raichle, 2015), this region appears to play a central role within this network, by exerting an excitatory effects on all the other nodes. On the other hand, in this experimental paradigm, the bilateral MTG served as the entry point for stimulus bound driving input; suggesting that sensory information is first processed in a region that is associated with language and narrative comprehension (Mar, 2011; Spreng et al., 2009).

Finally, in line with findings that support a key role of RTPJ in social cognition, particularly in self-other distinction (Aichhorn et al., 2006; Santiesteban, Banissy, et al., 2012; Rebecca Saxe & Kanwisher, 2003), our results showed that an increase in the RTPJ's sensitivity to afferent inputs from other nodes was associated with the process of focusing on the romantic partner (rather than on the self). Since this region has been causally involved in differentiating self and other representations (Santiesteban, Banissy, et al., 2012), a possible explanation is that in order to assume their partner's perspective, participants had to inhibit their own perspective (Steinbeis, 2015). However, the role of the RTPJ in the specific dynamics of (enhancing vs. inhibiting) self-other representations remains unclear.

The results of this study endorse the analysis of directed, effective connectivity, such as DCM, which can estimate the effect that one neural system exerts over another, to understand the dynamic interplay between the nodes of complex brain networks. Technically, this sort of analysis allows one to quantify intrinsic (self) connectivity that transpired to play a crucial role in this study. This is important because functional connectivity measures (such as those afforded by correlations or Granger causality) preclude such characterizations. This is particularly relevant when studying high-level psychological phenomena such as social cognition that entail different subprocesses and recruit distinct brain regions.

Regarding future directions, it would be interesting to extend the present analysis and use DCM to characterise context sensitive changes in connectivity within other social brain networks, (e.g. the empathy network). Another important contribution would be to study the dynamic interplay between different networks involved in social abilities. For example, to use DCM to characterise the coupling between more embodied or affective systems and more cognitive or conceptual systems. These analyses would allow us to better understand the functional integration of affective and cognitive aspects of social processing. Specifically, it would help establish if these respective networks are hierarchically related – in a way that mental state attribution depends on the capacity to share another's internal states. Finally, it would be interesting to examine the existence of similar effective connectivity patterns in other human dyads, such as parent-child or therapist-patient exchanges.

Acknowledgements and Author Contributions

This study was funded by BIAL Foundation (Grant number 87/12); by Portuguese Foundation for Science and Technology and the Portuguese Ministry of Education and Science through national funds and co-financed by FEDER through COMPETE2020 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653); by the postdoctoral scholarship UMINHO/BPD/18/2017 and by the FCT PhD scholarship (PD/BD/105963/2014). KJF is funded by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z).

JC and OFG designed the experimental design. JC and P.O-S collected fMRI data. SE performed the data analysis, interpretation and wrote the paper under the supervision of JC, JMS and KJF. All authors reviewed and approved the final draft.

CHAPTER V

DISCUSSION AND CONCLUSIONS

"From womb to tomb we are bound to others, past and present, and by each crime and every kind interaction we birth our future." (Arndt, Boden, Tykwer, Wachowski, & Wachowski, 2012). The success in a social environment is highly dependent on one's ability to flexibly understand the emotional and mental states of others. This capacity to understand other's inner states enables is what enables individuals to adjust their behaviour to the demands of the encounter. In the last decades, social neuroscience research has greatly advanced our understanding of different dimensions of social processing and of its neural basis (e.g. Carrington & Bailey, 2009; Farrow et al., 2001; Forbes & Grafman, 2010; Ochsner et al., 2004; Preston & Hofelich, 2012). The present dissertation intended to contribute to the further advance in this field.

Two main assumptions guided the research efforts of this work: the need for the use of analytic methods that assess the causal relationships among the regions that compose the already known social brain networks; the need for experimental tasks that can emulate more closely naturally occurring social situations. Departing from these assumptions this dissertation aimed at unravelling the neural basis (and respective functional integration) of social cognition in a dyadic context with a romantic partner using an effective connectivity approach. In particular, DCM was used to assess the internode communication within the social brain networks under study, in combination with fMRI resting state and an idiosyncratic social interaction paradigm intended to help overcome the lack of ecological validity of most previous studies found in the literature.

Three studies were conducted: study 1 focused on brain activity at rest, in particular on the relationship between DMN functional integration and perceived social ability; study 2 assessed the social processing (i.e. self and other processing) in a dyadic context with an intimate other using an fMRI paradigm derived from a previous real couple´s interaction; study 3 investigated the functional integration within the social brain, particularly the specific patterns of internode communication when making inferences about the other versus the self.

In study 1 we investigated the relationship between DMN's functional integration and social cognition. This study was motivated by the clear anatomical and functional overlap reported in the literature between the key regions of the DMN - PCC, mPFC and bilateral IPL (including TPJ) - and areas traditionally associated with social processing (e.g. Li et al., 2014; Mars et al., 2012; Schilbach et al., 2008; Schilbach et al., 2012). Based on the well-known relationship between DMN's activity and several processes crucial for social cognition

(Cox et al., 2012; Oliveira Silva et al., 2018; Otti et al., 2010; Vemuri & Surampudi, 2015), we aimed to explore whether the functional architecture of the DMN changed with the increase or decrease in self-perceived social cognition. Essentially, the goal of study 1 was to extend the findings of most previous studies that have used functional connectivity approaches to assess the relationship between DMN and social cognition (e.g. Oliveira Silva et al., 2018; Otti et al., 2010), by addressing DMN´s functional integration in terms of effective connectivity.

The DCM results showed that with the increase in self- perceived empathic abilities the functional architecture of the DMN changed with all nodes becoming simultaneously more self-sufficient and more susceptible to other nodes' influence. In addition, the PCC became the central node within the DMN by increasing its influence through an excitatory coupling on both IPLs, which propagated via the rIPL to the mPFC. These findings not only confirmed that there is a clear direct relationship between social cognition and the brain's intrinsic functional dynamics, but also illuminated the specific contribution of the PCC and the right IPL in DMN's functional integration during social processing. The result of the DMN's lateralization via the right IPL, is in line with data coming from lesion studies showing that empathic ability was compromised in patients with lesions in the right IPL, but not in patients with lesions in the left IPL (Shamay-Tsoory, Tomer, Goldsher, Berger, & Aharon-Peretz, 2004).

After this study in which we looked at individuals´ brain activity at rest in one social brain network, we intended to look at individuals´ brain activity while performing a social cognitive task. In other words, instead of measuring resting state brain activity and self-perceived social ability, we were now interested in the individuals´ task evoked neural activity during their performance in a social task. Thus, we used an fMRI task where each member of the couple watched a set of video-vignettes extracted from a previous real interaction with his/her partner and was asked to focus either on what he/she were feeling or on what the partner was feeling during the video. Results showed that despite the very significant overlap between the neural basis of self and other processing, self- inference engaged to a further extent brain regions traditionally linked with interoceptive processing and affect sharing, whereas other-based inference recruited to a greater extent brain areas involved in emotion recognition and TOM, in particular the rTPJ. Particularly the rTPJ is associated with the ability to distinguish and shift between self and other perspectives, which suggests that in order to assume the perspective of their romantic partner participants need to inhibit their own perspective (Steinbeis, 2015).

Having identified the brain activation maps for self and other processing, particularly which brain regions were commonly and specifically recruited in each condition, the next step was to investigate the causal information flow between these regions during this task. Thus, in study 3 we used DCM to characterize the coupling among the previously identified brain regions and see how this coupling changed when making inferences about the other, rather than the self. Specifically, we used DCM to characterize the role of each node within a well-known socio-cognitive network, constituted by the bilateral TPJ, the PCC/ prec and the bilateral MTG. DCM results have showed that, in this experimental paradigm, the bilateral MTG served as the entry point for stimulus related information, pointing to an initial processing of sensory information in this region known for language and narrative comprehension. In addition, by increasing its sensitivity to all its afferents, the RTPJ's played a decisive role in this network's functional integration, when making inferences about others, relative to self. Finally, the PCC played a central coordination role within this particular network, exerting an excitatory influence in all the other nodes, but receiving information only from the LMTG.

Altogether, the set of studies presented in this dissertation showed the clear direct relationship between the DMN's functional architecture and social cognition and uncovered the specific and crucial contribution that the right IPL, the PCC, the MTG and the mPFC have in social processing. Particularly, the right TPJ, played an important role in social brain networks' functional integration, which was demonstrated both when studying brain activity at rest and during the performance of a social task. Regarding the resting state results the rTPJ was a crucial node for the observed changes in DMN architecture as a function of a measure that assesses other-oriented feelings of concern and the tendency to adopt other's psychological point of view. Likewise, during the performance of the social task the rTPJ was the only node showing a change in coupling associated with focusing on the other's experience.

Furthermore, the findings presented in this dissertation provided further support for the close relationship between the brain activation associated with making inferences about the self and others, namely close others. This is in line with previous findings showing that closeness and intimacy is associated with the creation of overlapping brain structures between the self and the partner (Aron et al., 1992; Cheng et al., 2010; Péloquin & Lafontaine, 2010;

Waldinger et al., 2004), an overlap that seems to be related to the individuals' subjective reports of closeness (Cheng et al., 2010).

Another interesting overlap was found between the key regions of the DMN (study 1) – i.e. the PCC, mPFC and the bilateral IPLs (that comprise the TPJ) - and the regions recruited in the social cognition task (studies 2 and 3) – e.g. the bilateral MTG, the PCC and the bilateral TPJ. In fact this overlap was also evinced by the PCC exhibiting a similar excitatory orchestrator role within the DMN and the social cognition network selected; coordination role that has been consistently reported in the literature both within and between network (e.g. Deshpande, Santhanam, & Hu, 2011; Hagmann et al., 2008; van den Heuvel & Sporns, 2011). These results constitute further evidence supporting the consistent anatomical and functional overlap between the DMN and social cognition.

Finally, the results of study 1 and 3 in this dissertation illustrate in a very evident manner how the use of EC analytic methods added a more comprehensive level of explanation to the research questions under analysis. In particular, in study 1 the functional connectivity analysis – ICA – proposed that the mPFC was the only node associated with higher self-perceived empathy, whereas DCM revealed that this activation in the mPFC was due to systematic differences in coupling between the PCC and the rIPL. Moreover the results of the 3rd study emphasized the importance of assessing the connectivity between the nodes and within the node, causal characterizations that cannot be revealed using functional connectivity methods (Daunizeau et al., 2011; Friston, 1994, 2011b; Stephan & Friston, 2010).

Lastly in our view, by developing an fMRI task in which subjects were exposed to real videos of their romantic partner extracted from a previous interaction (studies 2 and 3), the present work made also an important methodological contribution. The creation of this type of task in which participants are exposed to a set of idiosyncratic stimuli without sacrificing experimental control constitutes a step forward in terms of ecological validity when compared to most existent fMRI tasks in which subjects are exposed to imagined or fictional targets. Despite the good tradeoff between internal and external validity of our task, this is still a relatively artificial situation. For example, each video vignette had only 20 seconds of duration, which is due to fMRI methodological constrains such as the limited duration of each trial. On the other hand, this paradigm is only a *proxy* of what would be an online couple´s interaction. In the future it would be interesting to investigate the neural bases of self and other processing during a real-time social interaction, as proposed by some authors who argue that social cognition is different when we are interacting or when we are observing others (Schilbach, 2014; Schilbach et al., 2013).

Furthermore, it is important to note that in the 3rd study's DCM analysis, we do not find the source of the neuromodulatory effects mediating the social process under study; we only identify the brain regions that constitute the targets of the context-sensitive modulation. This has to do with the fact that condition specific effects were modelled as experimental inputs – as opposed to an input from another part of the brain. In addition, in study 3 we only identify the brain regions in the selected network that first received the stimuli sensory information, having no information regarding the full pathway where the stimuli's information acts. It would be interesting to investigate how the sensory information moves from the primary visual and auditory areas until the MTG, thus confirming if the information really comes directly from the visual-sensory social networks (Alcalá-López et al., 2017) or if there is an intermediate network such as the affective network.

Besides, in this same study we only address the communication within a specific set of social cognitive regions. This happened because DCM only explains the functional integration within regions specifically recruited by the process or experimental task under analysis, i.e. it is task dependent. As such, since our paradigm did not recruit some socio-cognition key regions like the mPFC, we could not have a broader picture of how socio-cognition functional integration occurs. It would be of great interest to apply DCM to an fMRI paradigm that also recruited the mPFC, in order to clarify if in the process of social inference the right TPJ is indeed a receptor (study 3), or if similarly to what happens in the DMN (study 1), the right TPJ is just a mediator between the bilateral MTG/ PCC and higher regions of social processing, such as mPFC.

Lastly, considering that the affective dimensions emerged much earlier in the brain's evolution than higher cognitive capacities (Decety & Svetlova, 2012), and that our findings suggest an interaction between the affective and cognitive dimensions of social cognition (study 2), it would be interesting to use DCM to characterise the dynamic interplay between these two systems when making inferences about self and other's. This would help us clarify if in accordance with the premise of "simulation" theories of social cognition (ref) these networks are hierarchically related in a way that the high-level conceptual system depend on bottom up or embodied systems. In addition, it would be interesting to build on the results of study 1 which were focused on the DMN a cognitive and “mentalizing” network, and extend this analysis to networks known to be engaged in affective social processes such as the Salience Network, which is commonly recruited in affect sharing and empathy (Braun et al., 2012; Cao et al., 2014).

Overall, by bringing to light the neural mechanisms underlying social interactions, the findings presented in this dissertation are fundamental to social neuroscience research in itself. In particular this dissertation aimed to contribute to a research agenda that assumes that social neuroscience must increase our knowledge about real daily life social processes. By contributing to a better comprehension of the close relationship between self and other processing, in particular in dyadic context with an intimate other, this work has clear implications not only for couples functioning, but also for other human dyads such as mother-infant or helping professional relationships like physician-patient. In fact the comprehension of how social brain systems are modulated by real social interactions has crucial implications for the quality of all human´s relationships, from marital to mother-infant or labor relationships.

Furthermore, these results may also be important for translational social neuroscience approaches which aim at understanding and improving therapeutic interventions focused in the social deficits associated with different clinical disorders (e.g., Crespi & Badcock, 2008; Linden, 2006; Meyer-Lindenberg & Tost, 2012). In particular, clinical disorders in which the mechanisms of self and other processing are altered, such as affective disorders or autism (e.g. Dziobek et al., 2011; Schulte-Rüther et al., 2011). One future step of this line of research could be to test if by intervening in a self-oriented process we can increase or decrease other-oriented processing. In other words, if by cultivating the ability to read and regulate our own emotions, we will also improve the capacity to understand others and relate with them in more satisfactory ways.

REFERENCES

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, (12), 169–177. https://doi.org/10.1016/S0959-4388(02)00301-X

Aertsen, A., & Preibl, H. (1991). Dynamics of activity and connectivity in physiological neural networks. In Schuster H G (Ed.), *Nonlinear Dynamics and Neuronal Networks* (p. pp 281-301). Weinheim: VCH Verlag.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*(3), 1059–1068. https://doi.org/10.1016/j.neuroimage.2005.10.026

Alcalá-López, D., Smallwood, J., Jefferies, E., Van Overwalle, F., Vogeley, K., Mars, R. B., … Bzdok, D. (2017). Computing the social brain connectome across systems and states. *Cerebral Cortex*, *28*(7), 2207–2232. https://doi.org/10.1093/cercor/bhx121

Allen, P., Stephan, K. E., Mechelli, A., Day, F., Ward, N., Dalton, J., … McGuire, P. (2010). Cingulate activity and fronto-temporal connectivity in people with prodromal signs of psychosis. *NeuroImage*, *49*(1), 947–955. https://doi.org/10.1016/j.neuroimage.2009.08.038

Amft, M., Fox, P. T., Bzdok, D., Schilbach, L., Eickhoff, S. B., & Laird, A. R. (2014). *Definition and characterization of an extended social-affective default network. Brain Structure and Function* (Vol. 220). https://doi.org/10.1007/s00429-013-0698-0

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci*, *7*(4), 268–277. https://doi.org/10.1038/nrn1884

Arndt, S. , Boden, A. (Producers), Tykwer, T., Wachowski, L., & Wachowski, L. (Directors) (2012) Cloud Atlas [Motion Picture]. Babelsberg Studio, Germany

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology*, *60*(2), 241–253. https://doi.org/10.4324/9780203311851

Ashburner, J., & Friston, K. J. (2000). Voxel-Based Morphometry—The Methods. *NeuroImage*, *11*, 805–821. https://doi.org/10.1006/nimg.2000.0582

Baetens, K., Ma, N., Steen, J., & Van Overwalle, F. (2013). Involvement of the mentalizing network in social and non-social high construal. *Social Cognitive and Affective Neuroscience*, *9*(6), 817–824. https://doi.org/10.1093/scan/nst048

Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. (1992). Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, *25*(2), 390–397.

Bastos-Leite, A. J., Ridgway, G. R., Silveira, C., Norton, A., Reis, S., & Friston, K. J. (2015). Dysconnectivity within the default mode in first-episode schizophrenia: A stochastic dynamic causal modeling study with functional magnetic resonance imaging. *Schizophrenia Bulletin*, *41*(1), 144–153.

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and Empathy: Two Qualitatively Distinct Vicarious Emotions with Different Motivational Consequences. *Journal of Personality*, *55*(1), 19–39. https://doi.org/10.1111/j.1467-6494.1987.tb00426.x

Batson, C. D., Lishner, D. A., Cook, J., & Sawyer, S. (2005). Similarity and Nurturance: Two Possible Sources of Empathy for Strangers. *Basic and Applied Social Psychology*, *27*(1), 15–25. https://doi.org/10.1207/s15324834basp2701

Beach, S. R. H., Fincham, F. D., & Katz, J. (1998). Marital therapy in the treatment of depression: Toward a third generation of therapy and research. *Clinical Psychology Review*. https://doi.org/10.1016/S0272-7358(98)00023-3

Berthoz, S., Armony, J. L., Blair, R. J. R., & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain : A Journal of Neurology*, *125*(Pt 8), 1696–1708

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex December*, *19*, 2767–2796. https://doi.org/10.1093/cercor/bhp055

Böckler, A., Herrmann, L., Trautwein, F.-M., Holmes, T., & Singer, T. (2017). Know Thy Selves: Learning to Understand Oneself Increases the Ability to Understand Others. *Journal of Cognitive Enhancement*, *1*(2), 197–209. https://doi.org/10.1007/s41465-017-0023-6

Bookwala, J. (2005). The Role of Marital Quality in Physical Health During the Mature Years. *Journal of Aging and Health*, *17*(1), 85–104. https://doi.org/10.1177/0898264304272794

Botvinick, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E., & Prkachin, K. M. (2005).

Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *NeuroImage*, (25), 312–315. https://doi.org/10.1016/j.neuroimage.2004.11.043

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22.

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1528), 2359–2367. https://doi.org/10.1098/rstb.2009.0066

Braun, U., Plichta, M. M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., … Meyer-Lindenberg, A. (2012). Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *NeuroImage*, *59*(2), 1404–1412.

Breakspear, M. (2004). "Dynamic" Connectivity in Neural Systems. *Neuroinformatics*, *2*(2), 205–226. https://doi.org/10.1385/NI:2:2:205

Broca, P. (1865). Sur le siège de la faculté du langage articulé. *Bulletins de La Société d'anthropologie de Paris*, *6*(1), 377–393. https://doi.org/10.3406/bmsap.1865.9495

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience*, *6*(May), 1–19. https://doi.org/10.3389/fnhum.2012.00147

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Bugnyar, T., Reber, S. A., & Buckner, C. (2016). Ravens attribute visual access to unseen competitors. *Nature Communications*, *7*(1), 10506. https://doi.org/10.1038/ncomms10506

Buxton, R. B., & Frank, L. R. (1997). *A Model for the Coupling Between Cerebral Blood Flow and Oxygen Metabolism During Neural Stimulation*. *Journal of Cerebral Blood Flow and Metabolism* (Vol. 17).

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796.

Calhoun, V. D., Adali, T., McGinty, V. B., Pekar, J. J., Watson, T. D., & Pearlson, G. D. (2001). fMRI Activation in a Visual-Perception Task: Network of Areas Detected Using the General Linear Model and Independent Components Analysis. *NeuroImage*, *14*(5), 1080–1088.

https://doi.org/10.1006/nimg.2001.0921

Cao, H., Plichta, M. M., Schäfer, A., Haddad, L., Grimm, O., Schneider, M., … Tost, H. (2014). Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. *NeuroImage*, *84*, 888–900.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., Luigi Lenzi, G., & Raichle, M. E. (2003). *Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas*.

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping*, *30*(8), 2313–2335. https://doi.org/10.1002/hbm.20671

Carvalho, S., Leite, J., Galdo-Álvarez, S., & Gonçalves, Ó. F. (2012). The emotional movie database (EMDB): A self-report and psychophysiological study. *Applied Psychophysiology Biofeedback*, *37*(4), 279–294. https://doi.org/10.1007/s10484-012-9201-6

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, *129*(3), 564–583. https://doi.org/10.1093/brain/awl004

Chang, L. J., Yarkoni, T., Khaw, M. W., & Sanfey, A. G. (2012). Decoding the role of the insula in human cognition: Functional parcellation and large-scale reverse inference. *Cerebral Cortex*, *23*(3), 739–749. https://doi.org/10.1093/cercor/bhs065

Chen, C., Martínez, R. M., & Cheng, Y. (2018). The Developmental Origins of the Social Brain: Empathy, Morality, and Justice. *Frontiers in Psychology*, *9*, 2584. https://doi.org/10.3389/fpsyg.2018.02584

Chen, P., Ye, E., Jin, X., Zhu, Y., & Wang, L. (2019). Association between Thalamocortical Functional Connectivity Abnormalities and Cognitive Deficits in Schizophrenia. *Scientific Reports*, *9*(1), 2952. https://doi.org/10.1038/s41598-019-39367-z

Cheng, Y., Chen, C., Lin, C. P., Chou, K. H., & Decety, J. (2010). Love hurts: An fMRI study. *NeuroImage*, *51*(2), 923–929. https://doi.org/10.1016/j.neuroimage.2010.02.047

Christoff, K., & Gordon, A. (2008). The role of spontaneous thought in human cognition. *Decision Making*.

Cialdini, R., Brown, S., Lewis, B., Luce, C., & Neuberg, S. (1997). Reinterpreting the Empathy-Altruism Relationship: When One Into One Equals Oneness. *Journal of Personality and Social*

Psychology, *73*(3), 481–494.

Cohen, D., & Strayer, J. (1996). Empathy in conduct-disordered and comparison youth. *Developmental Psychology*, *32*(6), 988–998. https://doi.org/10.1037/0012-1649.32.6.988

Coutinho, J., Beiramar, A., Silva, C., Lema, A., Lima, V., Grace, R., … Sampaio, A. (2016). Validity evidence of the Portuguese version of the Interpersonal Reactivity Index for Couples. *Revista Avaliação Psicológica*, *14*(3), 309–317. https://doi.org/10.15689/ap.2015.1403.02

Coutinho, J. F., Silva, P. O., & Decety, J. (2014). Neurosciences, empathy, and healthy interpersonal relationships: recent findings and implications for counseling psychology. *Journal of Counseling Psychology*, *61*(January 2016), 541–548. https://doi.org/10.1037/cou0000021

Coutinho, J., Oliveira-Silva, · P, Mesquita, · A R, Barbosa, · M, Perrone-Mcgovern, · K M, & Gonçalves, · O F. (2017). Psychophysiological Reactivity in Couples During a Marital Interaction Task, *42*, 335–346. https://doi.org/10.1007/s10484-017-9380-2

Coutinho, J., Oliveira-Silva, P., Fernandes, E., Gonçalves, O. F., Correia, D., Perrone Mc-Govern, K., & Tschacher, W. (2018). Psychophysiological Synchrony During Verbal Interaction in Romantic Relationships. *Family Process*. https://doi.org/10.1111/famp.12371

Cox, C. L., Uddin, L. Q., Di martino, A., Castellanos, F. X., Milham, M. P., & Kelly, C. (2012). The balance between feeling and knowing: Affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. *Social Cognitive and Affective Neuroscience*, *7*(6), 727–737. https://doi.org/10.1093/scan/nsr051

Coyne, J. C., Rohrbaugh, M. J., Shoham, V., Sonnega, J. S., Nicklas, J. M., & Cranford, J. A. (2001). Prognostic Importance of Marital Quality for Survival of Congestive Heart Failure. *The American Journal of Cardiology*, *88*.

Craig, A. D. (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology*, *13*(4), 500–505. https://doi.org/10.1016/S0959-4388(03)00090-4

Crespi, B., & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behavioral and Brain Sciences*, *31*(3), 241–261. https://doi.org/10.1017/S0140525X08004214

Cui, L.-B., Liu, J., Wang, L.-X., Li, C., Xi, Y.-B., Guo, F., … Lu, H. (2015). Anterior cingulate cortex-related connectivity in first-episode schizophrenia: a spectral dynamic causal modeling study with functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, *9*, 589. https://doi.org/10.3389/fnhum.2015.00589

Cummings, E. M., & Davies, P. T. (2002). Effects of marital conflict on children: recent advances and emerging themes in process-oriented research. *Journal of Child Psychology and Psychiatry* , *43*(1), 31–63.

Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., … Raichle, M. E. (2006). *Consistent resting-state networks across healthy subjects* (Vol. 103). Retrieved from www.pnas.orgcgidoi10.1073pnas.0601417103

Darlington, R. B., Sharon, L. W., & Walberg, H. J. (1973). Canonical Variate Analysis and Related Techniques. *Review of Educational Research*, *43*(4), 433–454. https://doi.org/10.2307/1170075

Daunizeau, J., David, O., & Stephan, K. E. (2011). Dynamic causal modelling: A critical review of the biophysical and statistical foundations. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2009.11.062

Davey, C. G., Pujol, J., & Harrison, B. J. (2016). Mapping the self in the brain's default mode network. *NeuroImage*, *132*, 390–397. https://doi.org/10.1016/j.neuroimage.2016.02.022

Davis, M. H. (1994). *Empathy: A social psychological approach*. *Social psychology series.*

de Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences*, *10*(10), 435–441. https://doi.org/10.1016/j.tics.2006.08.008

de Waal, F. B. M. (2011). What is an animal emotion? *Annals of the New York Academy of Sciences*, *1224*(1), 191–206. https://doi.org/10.1111/j.1749-6632.2010.05912.x

Decety, J. (2010). The Neurodevelopment of Empathy in Humans. *Dev Neurosci*, *32*, 257–267. https://doi.org/10.1159/000317771

Decety, J., & Jackson, P. (2004). The Functional Architecture of Human Empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 71–100. https://doi.org/10.1177/1534582304267187

Decety, J., & Lamm, C. (2006). Human Empathy Through the Lens of Social Neuroscience. *The Scientific World JOURNAL*, *6*, 1146–1163. https://doi.org/10.1100/tsw.2006.221

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *Neuroscientist*, *13*(6), 580–593. https://doi.org/10.1177/1073858407304654

Decety, J., Michalska, K. J., & Akitsuki, Y. (2008). Who caused the pain? An fMRI investigation of empathy and intentionality in children. *Neuropsychologia*, *46*(11), 2607–2614.

https://doi.org/10.1016/j.neuropsychologia.2008.05.026

Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience*, *2*(1), 1–24. https://doi.org/10.1016/j.dcn.2011.05.003

Deng, Y., Guo, R., Ding, G., & Peng, D. (2012). Top-down modulations from dorsal stream in lexical recognition: An effective connectivity fmri study. *PLoS ONE*, *7*(3), e33337. https://doi.org/10.1371/journal.pone.0033337

Deshpande, G., Santhanam, P., & Hu, X. (2011). Instantaneous and Causal Connectivity in Resting State Brain Networks Derived from Functional MRI Data Gopikrishna. *NeuroImage*, *15*(54), 1043–1052. https://doi.org/10.3174/ajnr.A1256.Functional

Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., … Milham, M. P. (2008). Functional Connectivity of Human Striatum: A Resting State fMRI Study. *Cerebral Cortex*, *18*(12), 2735–2747. https://doi.org/10.1093/cercor/bhn041

Di, X., & Biswal, B. B. (2014). Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging. *NeuroImage*, *86*, 53–59. https://doi.org/10.1016/j.neuroimage.2013.07.071

Dimaggio, G., Lysaker, P. H., Carcione, A., Nicolò, G., & Semerari, A. (2008). Know yourself and you shall know the other... to a certain extent: Multiple paths of influence of self-reflection on mindreading. *Consciousness and Cognition*, *17*(3), 778–789. https://doi.org/10.1016/j.concog.2008.02.005

Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology*, *43*(3), 261–274. https://doi.org/10.1037/0022-0167.43.3.261

Dziobek, I., Prei??ler, S., Grozdanovic, Z., Heuser, I., Heekeren, H. R., & Roepke, S. (2011). Neuronal correlates of altered empathy and social cognition in borderline personality disorder. *NeuroImage*, *57*(2), 539–548. https://doi.org/10.1016/j.neuroimage.2011.05.005

Eisenberg, N., & Strayer, J. (1987). *Empathy and Its Development*. New York: Cambridge University Press.

Eres, R., Decety, J., Louis, W. R., & Molenberghs, P. (2015). Individual differences in local gray matter density are associated with differences in affective and cognitive empathy. *NeuroImage*, *117*, 305–310. https://doi.org/10.1016/j.neuroimage.2015.05.038

Esménio, S., Soares, J. M., Oliveira-Silva, P., Gonçalves, Ó. F., Decety, J., & Coutinho, J. (2019). Brain circuits involved in understanding our own and other´s internal states in the context of Romantic Relationships. *Social Neuroscience*, 17470919.2019.1586758. https://doi.org/10.1080/17470919.2019.1586758

Esménio, S., Soares, J. M., Oliveira-Silva, P., Zeidman, P., Razi, A., Gonçalves, Ó. F., … Coutinho, J. (2019). Using resting-state DMN effective connectivity to characterize the neurofunctional architecture of empathy. *Scientific Reports*, *9*(1), 2603. https://doi.org/10.1038/s41598-019-38801-6

Fairhall, S. L., Indovina, I., Driver, J., & MacAluso, E. (2009). The brain network underlying serial visual search: Comparing overt and covert spatial orienting, for activations and for effective connectivity. *Cerebral Cortex*, *19*(12), 2946–2958. https://doi.org/10.1093/cercor/bhp064

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews*, *35*(November 2015), 903–911. https://doi.org/10.1016/j.neubiorev.2010.10.009

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarrier, N., … Woodruff, P. W. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport*, *12*(11), 2433–2438. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11496124

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage*, *105*(December), 347–356. https://doi.org/10.1016/j.neuroimage.2014.10.043

Fiske, S. T., & Taylor, S. E. (2013). *Social Cognition: From Brains to Culture*. Sage.

Forbes, C. E., & Grafman, J. (2010). The Role of the Human Prefrontal Cortex in Social Cognition and Moral Judgment. *Annual Review of Neuroscience*, *33*(1), 299–324. https://doi.org/10.1146/annurev-neuro-060909-153230

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. *PLoS ONE*, *8*(1), 53286. https://doi.org/10.1371/journal.pone.0053286

Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, *68*(2), 113–143. https://doi.org/10.1016/S0301-0082(02)00076-X

Friston, K. (2009a). Causal Modelling and Brain Connectivity in Functional Magnetic Resonance

Imaging How Is the Brain Organised? Causality and Coupling How Can One Model Brain Connectivity? *PLoS Biology 7*(2), 1000033. https://doi.org/10.1371/journal

Friston, K. (2009b, February 17). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*. Public Library of Science. https://doi.org/10.1371/journal.pbio.1000033

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, *2*(1–2), 56–78. https://doi.org/10.1002/hbm.460020107

Friston, K. J. (2004). *Functional integration in the brain*. *Human Brain Function 2nd edn Academic Press*. San Diego.

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connectivity*, *1*(1), 13–36. https://doi.org/10.1089/brain.2011.0008

Friston, K. J., Harrison, L. M., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*, 1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7

Friston, K. J., Kahan, J., Biswal, B., & Razi, A. (2014). A DCM for resting state fMRI. *NeuroImage*, *94*, 396–407. https://doi.org/10.1016/j.neuroimage.2013.12.009

Friston, K., & Penny, W. (2011). Post hoc Bayesian model selection. *NeuroImage*, *56*(4), 2089–2099. https://doi.org/10.1016/j.neuroimage.2011.03.062

Frith, C. D. (1999). Interacting Minds-A Biological Basis. *Science*, *286*(5445), 1692–1695. https://doi.org/10.1126/science.286.5445.1692

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology*, *15*(17), R644–R645. https://doi.org/10.1016/j.cub.2005.08.041

Frith, C., & Frith, U. (2006). The Neural Basis of Mentalizing. *Neuron*, *50*(4), 531–534. https://doi.org/10.1016/j.neuron.2006.05.001

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. Lond. B*, (358), 459–473. https://doi.org/10.1098/rstb.2002.1218

Gabriel, B., Beach, S. R. H., & Bodenmann, G. (2010). Depression, Marital Satisfaction and Communication in Couples: Investigating Gender Differences. *Behavior Therapy*, *41*(3), 306–316. https://doi.org/10.1016/J.BETH.2009.09.001

Gallese, V. (2014). Bodily selves in relation: embodied simulation as second-person perspective on intersubjectivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130177–20130177. https://doi.org/10.1098/rstb.2013.0177

Gallese, V., & Goldman, A. (1998). Mirror neurons and the mind-reading. *Trens in Cognitive Sciences*, *2*(12), 493–501. https://doi.org/10.1016/S1364-6613(98)01262-5

Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences*, *15*(11), 512–519. https://doi.org/10.1016/j.tics.2011.09.003

Gallup, G. G., Frederick, M. J., & Nathan Pipitone, R. (2008). Morphology and Behavior: Phrenology Revisited. https://doi.org/10.1037/1089-2680.12.3.297

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, *20*(2), 226–241. https://doi.org/10.1016/j.cogbrainres.2004.02.012

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *Journal of Cognitive Neuroscience*, *18*(6), 932–948. https://doi.org/10.1162/jocn.2006.18.6.932

Goldman, A. (2006). *Simulating minds: the philosophy, psychology and neuroscience of mindreading*. Oxford University Press.

Gottman, J. M., & Levenson, R. W. (1988). The social psychophysiology of marriage. In P. Noller & M. A. Fitzpatrick (Ed.), *Monographs in social psychology of language, No. 1. Perspectives on marital interaction* (pp. 182–200). Clevedon, England: Multilingual Matters.

Greene, J. D. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, *293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Greene, S., & Griffin, W. (1998). Symptom study in context: effects of marital quality on signs of Parkinson's disease during patient-spouse interaction. *Psychiatry*, *61*(1), 35–45.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(1), 253–258. https://doi.org/10.1073/pnas.0135058100

Gu, X., Eilam-Stock, T., Zhou, T., Anagnostou, E., Kolevzon, A., Soorya, L., … Fan, J. (2015). Autonomic and brain responses associated with empathy deficits in autism spectrum disorder. *Human Brain Mapping*, *36*(9), 3323–3338. https://doi.org/10.1002/hbm.22840

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Van Wedeen, J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, *6*(7), 1479–1493. https://doi.org/10.1371/journal.pbio.0060159

Han, S., Fan, Y., Xu, X., Qin, J., Wu, B., Wang, X., … Mao, L. (2009). Empathic neural responses to others' pain are modulated by emotional contexts. *Human Brain Mapping*, *30*(10), 3227–3237. https://doi.org/10.1002/hbm.20742

Happé, F., Cook, J. L., & Bird, G. (2016). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. https://doi.org/10.1146/annurev-psych-010416-044046

Hoffman, M. L. (1984). Interaction of affect and cognition. In *Emotions, Cognition, and Behavior* (pp. 103–131).

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, *33*(3), 307–316. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4389335

Hohwy, J. (2007). Functional integration and the mind. *Synthese*, *159*(3), 315–328. https://doi.org/10.1007/s11229-007-9240-3

Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., & D'Esposito, M. (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive and Affective Neuroscience*, *3*(3), 204–217. https://doi.org/10.1093/scan/nsn019

Horwitz, B., Tagamets, M.-A., & McIntosh, A. R. (1999). Neural modeling, functional brain imaging, and cognition. *Trends in Cognitive Sciences*, *3*(3), 91–98. https://doi.org/10.1016/S1364-6613(99)01282-6

Huang, H., Nguyen, P. T., Schwab, N. A., Tanner, J. J., Price, C. C., & Ding, M. (2017). Mapping dorsal and ventral caudate in older adults: Method and validation. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10.3389/fnagi.2017.00091

Hyett, M. P., Breakspear, M. J., Friston, K. J., Guo, C. C., & Parker, G. B. (2015). Disrupted effective connectivity of cortical systems supporting attention and interoception in melancholia. *JAMA Psychiatry*, *72*(4), 350–358. https://doi.org/10.1001/jamapsychiatry.2014.2490

Hyett, M. P., Parker, G. B., Guo, C. C., Zalesky, A., Nguyen, V. T., Yuen, T., & Breakspear, M. (2015). Scene unseen: Disrupted neuronal adaptation in melancholia during emotional film viewing. *NeuroImage: Clinical*, *9*, 660–667. https://doi.org/10.1016/j.nicl.2015.10.011

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, *24*(3), 771–779. https://doi.org/10.1016/j.neuroimage.2004.09.006

Jirsa, V. K. (2004). Connectivity and Dynamics of Neural Information Processing. *Neuroinformatics*, *2*(2), 183–204. https://doi.org/10.1385/NI:2:2:183

Johnson, M. K., Raye, C. L., Mitchell, K. J., Touryan, S. R., Greene, E. J., & Nolen-Hoeksema, S. (2006). Dissociating medial frontal and posterior cingulate activity during self-reflection. *Social Cognitive and Affective Neuroscience*, *1*(1), 56–64. https://doi.org/10.1093/scan/nsl004

Kahan, J., & Foltynie, T. (2013). Understanding DCM: Ten simple rules for the clinician. *NeuroImage*, *83*, 542–549. https://doi.org/10.1016/j.neuroimage.2013.07.008

Kanske, P., Böckler, A., Trautwein, F.-M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and Theory of Mind. *NeuroImage*, *122*, 6–19.

Kawasaki, H., Naotsugu, T., Kovach, C. K., Nourski, K. V., Oya, H., Howard, M. A., & Adolphs, R. (2012). Processing of Facial Emotion in the Human Fusiform Gyrus. *J Cogn Neurosci*, *24*(6), 1358–1370. https://doi.org/10.1162/jocn

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., … Paulus, M. P. (2017). Review Interoception and Mental Health: A Roadmap. https://doi.org/10.1016/j.bpsc.2017.12.004

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding : an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2.Predictive

Kim et al. (2017). Altered Functional Connectivity of the Default Mode Network in Low-Empathy Subjects. *Yonsei Med J*, *58*(5), 1061–1065.

Kiran, S., Meier, E. L., Kapse, K. J., & Glynn, P. A. (2015). Changes in task-based effective connectivity in language networks following rehabilitation in post-stroke patients with aphasia. *Frontiers in Human Neuroscience*, *9*, 316. https://doi.org/10.3389/fnhum.2015.00316

Krahé, C., Springer, A., Weinman, J. A., & Fotopoulou, A. (2013). The Social Modulation of Pain: Others as Predictive Signals of Salience – a Systematic Review. *Frontiers in Human Neuroscience*, *7*(July), 1–21. https://doi.org/10.3389/fnhum.2013.00386

Lamm, C., Batson, C. D., & Decety, J. (2007). The Neural Substrate of Human Empathy : Effects of

Perspective-taking and Cognitive Appraisal. *Journal of Cognitive Neuroscience*, *19*(1), 42–58.

Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self–other representations in empathy: evidence from neurotypical function and socio-cognitive disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150083. https://doi.org/10.1098/rstb.2015.0083

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*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

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. *Journal of Cognitive Neuroscience*, *22*(2), 362–376. https://doi.org/10.1162/jocn.2009.21186

Lamm, C., Nausbaum, H. C., Meltzoff, A. N., & Decety, J. (2007). What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE*, *2*(12). https://doi.org/10.1371/journal.pone.0001292

Lang, S., Yu, T., Markl, A., Müller, F., & Kotchoubey, B. (2011). Hearing others' pain: Neural activity related to empathy. *Cognitive, Affective and Behavioral Neuroscience*, *11*(3), 386–395. https://doi.org/10.3758/s13415-011-0035-0

Lehnart, J., Neyer, F. J., & Eccles, J. (2010). Long-Term Effects of Social Investment: The Case of Partnering in Young Adulthood. *Journal of Personality*, *78*(2), 639–670. https://doi.org/10.1111/j.1467-6494.2010.00629.x

Li, W., Mai, X., & Liu, C. (2014). The default mode network and social understanding of others: what do brain connectivity studies tell us. *Frontiers in Human Neuroscience*, *8*, 74. https://doi.org/10.3389/fnhum.2014.00074

Linden, D. E. J. (2006). How psychotherapy changes the brain - The contribution of functional neuroimaging. *Molecular Psychiatry*, *11*(6), 528–538. https://doi.org/10.1038/sj.mp.4001816

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Wheelwright, S. J., Sadek, S. A., Suckling, J., & Baron-Cohen, S. (2009). Shared Neural Circuits for Mentalizing about the Self and Others. *Journal of Cognitive Neuroscience*, *22*(7), 1623–1635. https://doi.org/10.1162/jocn.2009.21287

Long, E. C. J. (1990). Measuring Dyadic Perspective-Taking: Two Scales for Assessing Perspective-

Taking in Marriage and Similar Dyads. *Educational and Psychological Measurement*, *50*(1), 91–103. https://doi.org/10.1177/0013164490501008

Lu, H., Zou, Q., Gu, H., Raichle, M. E., Stein, E. A., & Yang, Y. (2012). Rat brains also have a default mode network. *Proceedings of the National Academy of Sciences*, *109*(10), 3979–3984. https://doi.org/10.1073/pnas.1200506109

Mar, R. A. (2011). The Neural Bases of Social Cognition and Story Comprehension. *Annual Review of Psychology*, *62*(1), 103–134. https://doi.org/10.1146/annurev-psych-120709-145406

Mars, R. B., Neubert, F.-X., Noonan, M. P., Sallet, J., Toni, I., & Rushworth, M. F. S. (2012). On the relationship between the "default mode network" and the "social brain." *Frontiers in Human Neuroscience*, *6*(June), 1–9. https://doi.org/10.3389/fnhum.2012.00189

Mayne, T. J., O'leary, A., McCrady, B., Contrada, R., & Labouvie, E. (1997). The differential effects of acute marital distress on emotional, physiological and immune functions in maritally distressed men and women. *Psychology & Health*, *12*(2), 277–288.

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*, *214*(5–6), 655–667. https://doi.org/10.1007/s00429-010-0262-0.Saliency

Meyer-Lindenberg, A., & Tost, H. (2012). Neural mechanisms of social risk for psychiatric disorders. *Nature Neuroscience*, *15*(5), 663–668. https://doi.org/10.1038/nn.3083

Meyer, M. L., Masten, C. L., Ma, Y., Wang, C., Shi, Z., Eisenberger, N. I., & Han, S. (2013). Empathy for the social suffering of friends and strangers recruits distinct patterns of brain activation. *Social Cognitive and Affective Neuroscience*, *8*(4), 446–454. https://doi.org/10.1093/scan/nss019

Minkova, L., Scheller, E., Peter, J., Abdulkadir, A., Kaller, C. P., Roos, R. A., … Klöppel, S. (2015). Detection of Motor Changes in Huntington's Disease Using Dynamic Causal Modeling. *Frontiers in Human Neuroscience*, *9* (November), 1–13.

Mitchell, J. P. (2009). Inferences about mental states. Philosofical Transactions of the Royal Society B: Biological Science, 364 (1521), 1309–1316. https://doi.org/10.1098/rstb.2008.0318

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron*, *50*(4), 655–663. https://doi.org/10.1016/j.neuron.2006.03.040

Mkael, S., Horan, C. H., Leite, I. M., Buckley, C., Irani, S. R., Stephan, K. E., … Moran, R. J. (2018). Ion channels in EEG: isolating channel dysfunction in NMDA receptor antibody encephalitis. *Brain*, *141*(1691), 1691–1702. https://doi.org/10.1093/brain/awy136

Moran, R. J., Stephan, K. E., Dolan, R. J., & Friston, K. J. (2011). Consistent spectral predictors for dynamic causal models of steady-state responses. *NeuroImage*, *55*(4), 1694–1708. https://doi.org/10.1016/j.neuroimage.2011.01.012

Moran, R., Pinotsis, D. A., Friston, K., Robinson, P., & Roberts, J. (2013). Neural masses and fields in dynamic causal modeling. https://doi.org/10.3389/fncom.2013.00057

Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain-A meta-analysis of imaging studies on the self. https://doi.org/10.1016/j.neuroimage.2005.12.002

Novembre, G., Zanon, M., & Silani, G. (2015). Empathy for social exclusion involves the sensory-discriminative component of pain: a within-subject fMRI study. *Social Cognitive and Affective Neuroscience*, *10*(2), 153–164. https://doi.org/10.1093/scan/nsu038

Ochsner, K. N., Knierim, K., Ludlow, D. H., Hanelin, J., Ramachandran, T., Glover, G., & Mackey, S. C. (2004). *Reflecting upon Feelings: An fMRI Study of Neural Systems Supporting the Attribution of Emotion to Self and Other*.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, *89*(13), 5951–5955. https://doi.org/10.1073/pnas.89.13.5951

Oliveira Silva, P., Maia, L., Coutinho, J., Frank, B., Soares, J. M., Sampaio, A., & Gonçalves, Ó. (2018). Empathy by default: Correlates in the brain at rest. *Psicothema*, *30*(1), 97–103. https://doi.org/10.7334/psicothema2016.366

Osnes, B., Hugdahl, K., & Specht, K. (2011). Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *NeuroImage*, *54*(3), 2437–2445. https://doi.org/10.1016/j.neuroimage.2010.09.078

Otti, A., Guendel, H., Läer, L., Wohlschlaeger, A. M., Lane, R. D., Decety, J., … Noll-Hussong, M. (2010). I know the pain you feel-how the human brain's default mode predicts our resonance to another's suffering. *Neuroscience*, *169*(1), 143–148. https://doi.org/10.1016/j.neuroscience.2010.04.072

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Péloquin, K., & Lafontaine, M.-F. (2010). Measuring empathy in couples: validity and reliability of the Interpersonal Reactivity Index for couples. *Journal of Personality Assessment*, *92*(June 2012), 146–157. https://doi.org/10.1080/00223890903510399

Perry, D., Hendler, T., & Shamay-Tsoory, S. G. (2012). Can we share the joy of others? Empathic neural responses to distress vs joy. *Social Cognitive and Affective Neuroscience*, *7*(8), 909–916. https://doi.org/10.1093/scan/nsr073

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(01), 1–71. https://doi.org/10.1017/S0140525X02000018

Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review*, *4*(1), 24–33. https://doi.org/10.1177/1754073911421378

Price, C. J., & Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology*, *22*(3–4), 262–275. https://doi.org/10.1080/02643290442000095

Qin, P., & Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage*, *57*(3), 1221–1233. https://doi.org/10.1016/j.neuroimage.2011.05.028

Raichle, M. E. (2015). The Brain's Default Mode Network. *Annual Review of Neuroscience*, *38*(1), 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Ramsey, R., Hansen, P., Apperly, I., & Samson, D. (2013). Seeing It My Way or Your Way: Frontoparietal Brain Areas Sustain Viewpoint-independent Perspective Selection Processes. *Journal of Cognitive Neuroscience*, *25*(5), 670–684. https://doi.org/10.1162/jocn_a_00345

Razi, A., Kahan, J., Rees, G., & Friston, K. J. (2015). Construct validation of a DCM for resting state fMRI. *NeuroImage*, *106*, 1–14. https://doi.org/10.1016/j.neuroimage.2014.11.027

Rehman, U. S., Gollan, J., & Mortimer, A. R. (2008). The marital context of depression: Research, limitations, and new directions. *Clinical Psychology Review*, *28*(2), 179–198. https://doi.org/10.1016/J.CPR.2007.04.007

Rehme, A. K., Eickhoff, S. B., Wang, L. E., Fink, G. R., & Grefkes, C. (2011). Dynamic causal modeling of cortical activity from the acute to the chronic stage after stroke. *NeuroImage*, *55*(3), 1147–1158. https://doi.org/10.1016/j.neuroimage.2011.01.014

Reniers, R. L. E. P., Völlm, B. A., Elliott, R., & Corcoran, R. (2014). Empathy, ToM, and self-other differentiation: An fMRI study of internal states. *Social Neuroscience*, *9*(1), 50–62. https://doi.org/10.1080/17470919.2013.861360

Robles, T. F., & Kiecolt-Glaser, J. K. (2003). The physiology of marriage: pathways to health. *Physiology & Behavior*, *79*, 409–416. https://doi.org/10.1016/S0031-9384(03)00160-4

Rogers, B. P., Morgan, V. L., Newton, A. T., & Gore, J. C. (2007). Assessing functional connectivity in the human brain by fMRI. *Magnetic Resonance Imaging*, *25*, 1347–1357. https://doi.org/10.1016/j.mri.2007.03.007

Rosa, M. J., Friston, K., & Penny, W. (2012). Post-hoc selection of dynamic causal models. *Journal of Neuroscience Methods*, *208*(1), 66–78. https://doi.org/10.1016/j.jneumeth.2012.04.013

Rutgen, M., Seidel, E.-M., Rie ansky, I., & Lamm, C. (2015). Reduction of Empathy for Pain by Placebo Analgesia Suggests Functional Equivalence of Empathy and First-Hand Emotion Experience. *Journal of Neuroscience*, *35*(23), 8938–8947.

Rytsar, R., Fornari, E., Frackowiak, R. S., Ghika, J. A., & Knyazeva, M. G. (2011). Inhibition in early Alzheimer’s disease: An fMRI-based study of effective connectivity. *NeuroImage*, *57*(3), 1131–1139. https://doi.org/10.1016/j.neuroimage.2011.05.029

Saarela, M. V., Hlushchuk, Y., Williams, A. C. d. C., Schurmann, M., Kalso, E., & Hari, R. (2006). The Compassionate Brain: Humans Detect Intensity of Pain from Another’s Face. *Cerebral Cortex*, *17*(1), 230–237. https://doi.org/10.1093/cercor/bhj141

Salley, B., & Colombo, J. (2016). Conceptualizing Social Attention in Developmental Research. *Social Development*, *25*(4), 687–703. https://doi.org/10.1111/sode.12174

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. *Current Biology*, *22*(23), 2274–2277. https://doi.org/10.1016/j.cub.2012.10.018

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, *122*(2), 228–235. https://doi.org/10.1016/j.cognition.2011.11.004

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding Other Minds: Linking Developmental Psychology and Functional Neuroimaging. *Annual Review of Psychology*, *55*(1), 87–124. https://doi.org/10.1146/annurev.psych.55.090902.142044

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, *1*(3), 229–234. https://doi.org/10.1093/scan/nsl034

Saxe, R., & Powell, L. J. (2006). Is it the thought that counts?Specific Brain Regions for One Component of Theory of Mind. *Psychological Science : A Journal of the American Psychological Society / APS*, *17*(8), 692–699. https://doi.org/10.1017/S0140525X00053607

Schilbach, L. (2014). On the relationship of online and offline social cognition. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00278

Schilbach, L., Bzdok, D., Timmermans, B., Fox, P. T., Laird, A. R., Vogeley, K., & Eickhoff, S. B. (2012). Introspective Minds: Using ALE meta-analyses to study commonalities in the neural correlates of emotional processing, social & unconstrained cognition. *PLoS ONE*, *7*(2), 30920. https://doi.org/10.1371/journal.pone.0030920

Schilbach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Fink, G. R., & Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Consciousness and Cognition*, *17*(2), 457–467.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience 1. https://doi.org/10.1017/S0140525X12000660

Schlösser, R. G. M., Wagner, G., Koch, K., Dahnke, R., Reichenbach, J. R., & Sauer, H. (2008). Fronto-cingulate effective connectivity in major depression: A study with fMRI and dynamic causal modeling. *NeuroImage*, *43*(3), 645–655.

Schulte-Rüther, M., Greimel, E., Markowitsch, H. J., Kamp-Becker, I., Remschmidt, H., Fink, G. R., & Piefke, M. (2011). Dysfunctions in brain networks supporting empathy: An fMRI study in adults with autism spectrum disorders. *Social Neuroscience*, *6*(1), 1–21. https://doi.org/10.1080/17470911003708032

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, *42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Shamay-Tsoory, S. G. (2011). The neural bases for empathy. *Neuroscientist*, *17*(1), 18–24. https://doi.org/10.1177/1073858410379268

Shamay-Tsoory, S., Tomer, R., Goldsher, D., Berger, B. D., & Aharon-Peretz, J. (2004). Impairment in cognitive and affective empathy in patients with brain lesions: anatomical and cognitive correlates. *J Clin Exp Neuropsychol*, *26*(8), 1113–1127.

Sharaev, M. G., Zavyalova, V. V, Ushakov, V. L., Kartashov, S. I., & Velichkovsky, B. M. (2016). Effective Connectivity within the Default Mode Network: Dynamic Causal Modeling of Resting-State fMRI Data. *Frontiers in Human Neuroscience*, *10*(February), 14.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*(July), 334–340. https://doi.org/10.1016/j.tics.2009.05.001

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, *1156*, 81–96. https://doi.org/10.1111/j.1749-6632.2009.04418.x

Singer, T., Seymour, B., O'Doherty, J., Dolan, R. J., Kaube, H., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science (New York, N.Y.)*, *303*(5661), 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*(7075), 466–469. https://doi.org/10.1038/nature04271

Soares, J., Magalhães, R., Moreira, P., Sousa, A., Ganz, E., Sampaio, A., … Sousa, N. (2016). A hitchhiker’s guide to functional Magnetic Resonance Imaging. *Frontiers in Neuroscience*, *10*(November), 515. https://doi.org/10.3389/FNINS.2016.00515

Sommer, M., Döhnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. https://doi.org/10.1016/j.neuroimage.2007.01.042

Sonty, S. P., Mesulam, M.-M., Weintraub, S., Johnson, N. A., Parrish, T. B., & Gitelman, D. R. (2007). Altered Effective Connectivity within the Language Network in Primary Progressive Aphasia. *Journal of Neuroscience*, *27*(6), 1334–1345.

Sowden, S., & Catmur, C. (2015). The Role of the Right Temporoparietal Junction in the Control of Imitation, (April), 1107–1113. https://doi.org/10.1093/cercor/bht306

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, *21*(3), 489–510. https://doi.org/10.1162/jocn.2008.21029

Steinbeis, N. (2015). The role of self – other distinction in understanding others ’ mental and emotional states : neurocognitive mechanisms in children and adults.

Stephan, K. E., & Friston, K. J. (2010). Analyzing effective connectivity with functional magnetic resonance imaging. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(3), 446–459. https://doi.org/10.1002/wcs.58

Tang, R., Razi, A., Friston, K. J., & Tang, Y.-Y. (2016). Mapping Smoking Addiction Using Effective Connectivity Analysis. *Frontiers in Human Neuroscience*, *10*(May), 1–9. https://doi.org/10.3389/fnhum.2016.00195

Tomasi, D., & Volkow, N. D. (2012). Resting functional connectivity of language networks: characterization and reproducibility. *Mol Psychiatry*, *17*(8), 841–854. https://doi.org/10.1038/mp.2011.177

Toyomaki, A., & Murohashi, H. (2013). “Salience network” dysfunction hypothesis in autism spectrum disorders. *Japanese Psychological Research*, *55*(2), 175–185. https://doi.org/10.1111/jpr.12012

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic

Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Uddin, L. Q. (2017). *Salience network of the human brain*. Academic Press.

van den Heuvel, M. P., & Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *Journal of Neuroscience*, *31*(44), 15775–15786. https://doi.org/10.1523/JNEUROSCI.3539-11.2011

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, *48*(3), 564–584. https://doi.org/10.1016/j.neuroimage.2009.06.009

Vemuri, K., & Surampudi, B. R. (2015). An exploratory investigation of functional network connectivity of empathy and default mode networks in a free-viewing task. *Brain Connectivity*, (502), 1–41.

Vogt, B. A., & Laureys, S. (2005). Posterior cingulate, precuneal and retrosplenial cortices: Cytology and components of the neural network correlates of consciousness. *Progress in Brain Research*. https://doi.org/10.1016/S0079-6123(05)50015-3

Völlm, B. A., Taylor, A. N. W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., … Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, *29*(1), 90–98. https://doi.org/10.1016/j.neuroimage.2005.07.022

Waldinger, R. J., Hauser, S. T., Schulz, M. S., Allen, J. P., & Crowell, J. A. (2004). Reading others emotions: The role of intuitive judgments in predicting marital satisfaction, quality, and stability. *Journal of Family Psychology : JFP : Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *18*(1), 58–71. https://doi.org/10.1037/0893-3200.18.1.58

Wernicke, C. (1874). Der aphasische Symptomenkomplex. In *Der aphasische Symptomenkomplex* (pp. 1–70). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-65950-8_1

Wharton, S. J., Basu, S. P., & Ashe, H. L. (2005). Rhesus Monkeys Attribute Perceptions to Others.

Current Biology, *15*, 447–452. https://doi.org/10.1016/j

Whitfield-Gabrieli, S., & Ford, J. M. (2012). Default Mode Network Activity and Connectivity in Psychopathology. *Annual Review of Clinical Psychology*, *8*, 49–76. https://doi.org/10.1146/annurev-clinpsy-032511-143049

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*(3), 655–664. https://doi.org/10.1016/S0896-6273(03)00679-2

Winter, K., Spengler, S., Bermpohl, F., Singer, T., & Kanske, P. (2017). Social cognition in aggressive offenders: Impaired empathy, but intact theory of mind. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-00745-0

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology*, *50*(2), 314–321. https://doi.org/10.1037/0022-3514.50.2.314

Yu, H., Chen, X., Liu, J., & Zhou, X. (2013). Gum Chewing Inhibits the Sensory Processing and the Propagation of Stress-Related Information in a Brain Network. *PLoS ONE*, *8*(4), 2–9. https://doi.org/10.1371/journal.pone.0057111

Zaki, J., & Ochsner, K. (2009). The need for a cognitive neuroscience of naturalistic social cognition. *Annals of the New York Academy of Sciences*, *1167*, 16–30. https://doi.org/10.1111/j.1749-6632.2009.04601.x

Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences*, *106*(27), 11382–11387. https://doi.org/10.1073/pnas.0902666106

Zeidman, P., Jafarian, A., Corbin, N., Seghier, M. L., Razi, A., Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. *NeuroImage*. https://doi.org/10.1016/J.NEUROIMAGE.2019.06.031

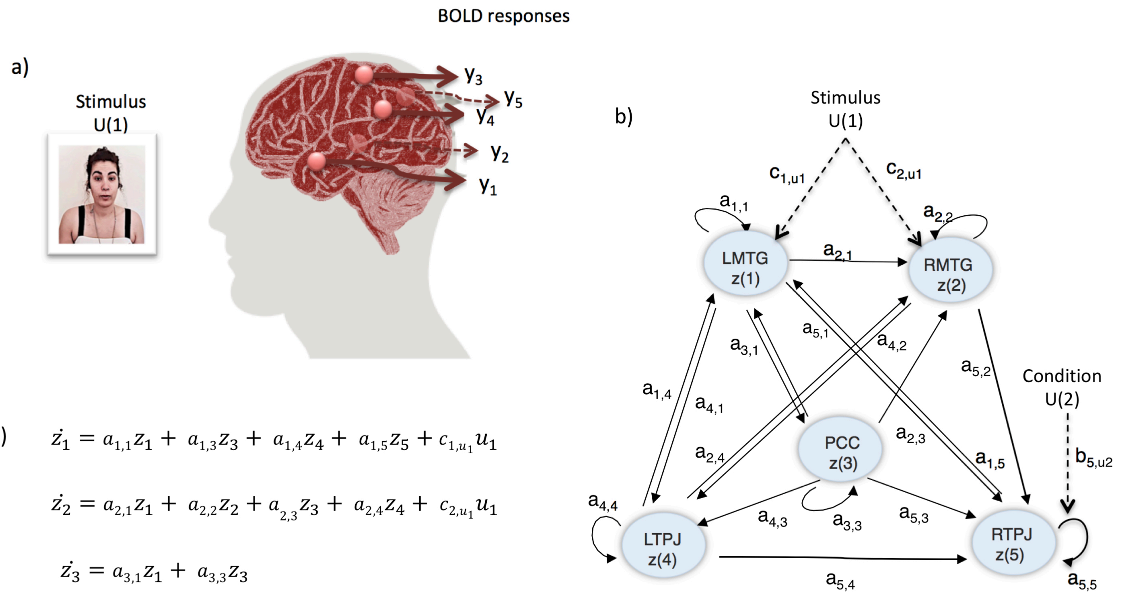
Zhang, H., Wei, X., Tao, H., Mwansisya, T. E., Pu, W., He, Z., … Xue, Z. (2013). Opposite Effective Connectivity in the Posterior Cingulate and Medial Prefrontal Cortex between First-Episode Schizophrenic Patients with Suicide Risk and Healthy Controls. *PLoS ONE*, *8*(5), 1–8. https://doi.org/10.1371/journal.pone.0063477

APPENDICES

Appendix A

Supplementary Material – Introduction

Supplementary Figure I -1 DCM fMRI experiment in study 3. Parameters $a_{k,j}$ represents the EC from node *j* to *k*; $c_{k,u}$ denote the effect of the driving input *u* on target node *k*; and $b_{k,u}$ denote the modulatory effect of the input *u* on target connection *k*. Adapted from Kahan & Foltynie (2013).



c)

$$\dot{z_1} = a_{1,1}z_1 + a_{1,3}z_3 + a_{1,4}z_4 + a_{1,5}z_5 + c_{1,u_1}u_1$$

$$\dot{z_2} = a_{2,1}z_1 + a_{2,2}z_2 + a_{2,3}z_3 + a_{2,4}z_4 + c_{2,u_1}u_1$$

$$\dot{z_3} = a_{3,1}z_1 + a_{3,3}z_3$$

$$\dot{z_4} = a_{4,1}z_1 + a_{4,2}z_2 + a_{4,3}z_3 + a_{4,4}z_4$$

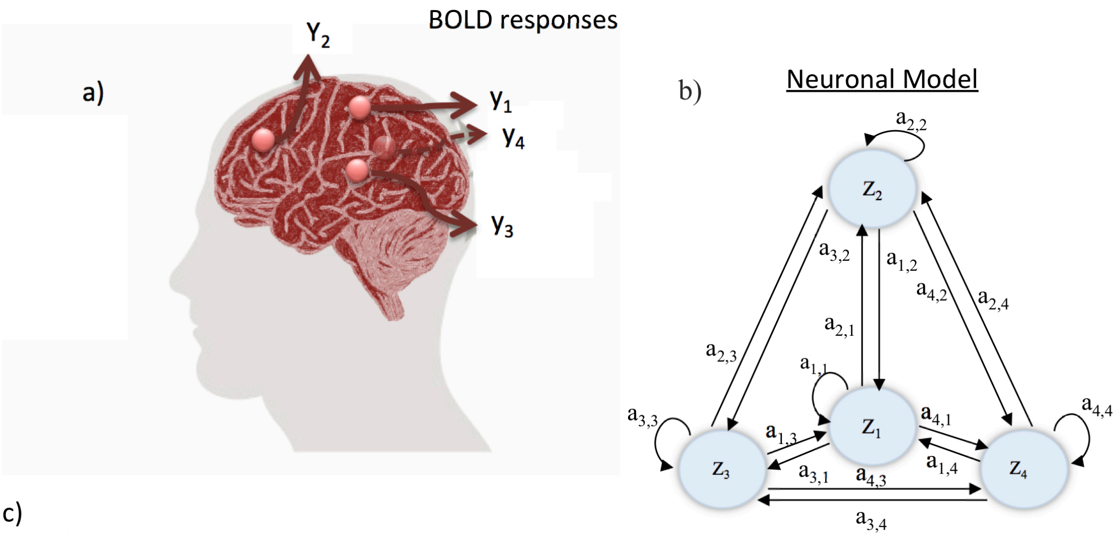
$$\dot{z_5} = a_{5,1}z_1 + a_{5,2}z_2 + a_{5,3}z_3 + a_{5,4}z_4 + a_{5,5}z_5 + u_2 z_5 b_{5,u_2}$$

In brief, this fMRI paradigm used a standard block design with a 2 x 3 factorial structure that included a self-other factor and three levels of valence. However, in the DCM analysis we focus only on the specific effect of the self-other factor. As such, the neuronal model consisted of one 'driving input' (u_1); u(1) = visual stimulation from videos-vignettes of romantic partner expressing personal emotional contents; and one modulatory effect (u_2); u(2) = the effect of focusing on the partner's internal states.

Five regions of interest were selected, namely lMTG, rMTG, PCC, lTPJ and r TPJ. The BOLD signal time series of these regions, represented by y_n was collected. Then the specific pattern of (hidden) neural activity z_n that underwrite these time series was computed.

Considering that the neural activity in region r ($\dot{z_r}$) is the sum of intrinsic regional connectivity and extrinsic or between regions connectivity, the regional self-inhibitory activity of region r is displayed by the value $a_{r,r}$; whereas the influence that the region r1 has on the region r2, is determine by $a_{r2,r1}$. In particular, in this experiment the driving input u_1 had an effective strength of c_{1,u_1} on the left and right MTG, which results on the addition of the terms $c_{1,u_1}u_1$ in $\dot{z_1}$ and $c_{2,u_1}u_1$ in $\dot{z_2}$. Lastly, the modulatory effect u_2 has an effective strength of b_{5,u_2} on the rTPJ intrinsic connection $a_{5,5}$; resulting on the addition of the term $u_2 z_5 b_{5,u_2}$ in $\dot{z_5}$. As one can see in the neural equations, task-based DCM adapt to the specificities of the paradigm simply by accommodating (any) driving or modulatory input.

Supplementary Figure I- 2 Scheme of the DCM resting state fMRI experiment conducted in study 1. Coupling parameters $a_{k,j}$ represents the EC from node *j* to *k*. Adapted from Kahan & Foltynie (2013).



c)

$$\dot{z_n} = a_{n,1}z_1 + a_{n,2}z_2 + a_{n,3}z_3 + a_{n,4}z_4 + v$$

Appendix B

Supplementary Material – Chapter IV

Supplementary Figure IV - 1. Results of specifying a modulatory effect in every intrinsic (self-) connection. (a) Parameters posterior estimates (EPs). (b) Connectivity matrix. (c) Structure and parameters. The black lines/values illustrate the (natural) connectivity between brain regions. The arrows in blue represent the modulatory effects. The numbers are the strength of connectivity (Hz).