Anchoring the self in the neural monitoring of visceral signals : how heartbeat-evoked responses encode the self

Mariana Babo-Rebelo

Thèse de Doctorat de l'Université Pierre et Marie Curie

Ecole Doctorale Cerveau Cognition Comportement

Présentée par Mariana Babo-Rebelo

Pour obtenir le grade de docteur de l'Université Pierre et Marie Curie

Anchoring the self in the neural monitoring of visceral signals: How heartbeat-evoked responses encode the self.

Thèse sous la direction de Dr. Catherine Tallon-Baudry

Laboratoire de Neurosciences Cognitives, UMR 960, DEC/ENS/Inserm

Thèse soutenue le 3 mai 2017 devant un jury composé de :

Prof. Hugo Critchley	Rapporteur
Dr. Antoine Lutz	Rapporteur
Prof. Patrick Haggard	Examinateur
Dr. Philippe Fossati	Examinateur
Dr. Virginie van Wassenhove	Examinateur
Dr. Catherine Tallon-Baudry	Directrice de thèse

Acknowledgments

I would like to start by thanking Catherine. Thank you for being a great supervisor, for introducing me to science, for teaching me rigor and perseverance. When I look back, I realize how much I learnt and how fulfilling it was for me to work with you. I remember feeling amazed during those meetings where you unraveled the mysteries of data. Thank you for your support and enthusiasm at all times, and for always pushing me forward.

I would also like to thank Prof. António Damásio, with whom I started working on the self, and who supported me for my grant application.

The MEG would not be the same without the wonderful people of the MEG center. Thank you Christophe Gitton, Denis Schwartz, Jean-Didier Lemarechal, Laurent Hugueville, Lydia Yahia Cherif and Nathalie George, for all your help and for creating such a pleasant atmosphere, even in this windowless underground. My warmest thanks go to Antoine Ducorps, whom I deeply admire and whose kindness, patience and generosity are absolutely unique.

Thanks also to all the dedicated people involved in the intracranial recordings, at the Pitié-Salpêtrière, from the technicians (Corinne, Joseph, Marie-Françoise…) to the doctors (Claude Adam, Vincent Navarro, Dominique Hasboun) and researchers (Katia Lehongre, Imen El Karoui).

The LNC has been since the beginning a very welcoming place to work. Thanks to all the PIs for maintaining this spirit, and thanks to the administrative staff (Marine and Laura) for being so efficient, helpful and kind.

What can I say about the wonderful LNC crew? Here I found more than just colleagues. Marion (vive les madeleines au chocolat!), Marwa (my conference buddy!), Emma (A vos ordres, chef Vilarem!), Michele (you are always cheerful, and thanks for proof-reading this manuscript), Hannah (memories from Greece, cantine and yoga… all great!), Margaux (thanks for keeping everyone hydrated!) – thank you all. Christina, you are just awesome, and thank you for your pink ballet elephants in the moments I most needed them. Thanks for being energetic and enthusiastic, you made my days brighter.

I would like also to thank everyone who has been part of the Visual Cognition Group, for creating such a friendly and collaborative atmosphere: Anne B., Anne U., Clémence, Diego, Heyong, Jeroen, Max, Nick, Nicolai, Samuel, Victor. Special thanks to Kayeon (you are the cutest), Ignacio (you are an explosion of good vibes), Craig (for all brainstorming sessions) and Stephen (what is Stephen-proof is reviewer-proof!). Florence, you are the personification of empathy. Thank you for being such a good friend, thank you for all the endless conversations about any topic and for being so sincerely supportive. My deepest thanks go to Damiano. You are irreplaceable. Thank you for being my most faithful pillar, for believing in this and for making me laugh as easily as you can light a match.

Finally, a special word to my family who has always been supportive and curious about my work. Maria João, Rui André, Pedro and Filipa, thank you for always keeping a caring eye on your little sister. Avô Manel, I wish I had told you I was doing a PhD; I can vividly picture the face of profound joy you would make now that I am about to finish it. Avó Lena, you fell into a state of semi-consciousness, the month I started this project. I wish I was able to use this neuroscience to help you. And finally my parents, I am so grateful to you, for all your care and love. You made Lisbon a city next to Paris, by always closely accompanying me, such that all my ups and downs became yours. Above all, I thank you because together and in your unique ways you made me who I am.

Obrigada.

This thesis is dedicated to Sophie, who passed away before she could fulfill her dream of becoming a scientist.

From *Le Coeur, socle du soi*
Cerveau & Psycho, March 2017

Abstract

The self has been hypothesized to be anchored in the neural monitoring of visceral signals; yet experimental evidence is still scarce. The main goal of this thesis was to directly address this question, by experimentally testing whether we could find a link between heart-brain coupling and the self. We operationalized the concept of self by defining two self-dimensions: the experiential "I" and the introspective "Me".

We showed in a first magnetoencephalography (MEG) experiment, that the self-relatedness of spontaneous thoughts was encoded in the amplitude of heartbeat-evoked responses (HERs) occurring in midline regions of the default-network. More precisely, we found that HERs in the posterior cingulate cortex / ventral precuneus encoded the "I" dimension, whereas the "Me" dimension was associated with HERs in the ventromedial prefrontal cortex. We additionally demonstrated that these results were specific to each self-dimension, thereby supporting a biological dissociation between the "I" and the "Me".

In a second study, we replicated and extended these results using intracranial recordings of epileptic patients and new analyses of the MEG data. Here, HER amplitude co-varied with the self-relatedness of spontaneous thoughts, at the single trial level. Moreover, a region of interest analysis of the right anterior insula showed that HERs in this region were also associated with the "I".

A third study (in prep.) aimed at testing these results in the context of oriented thoughts, in an imagination task. We found that HER amplitude in medial motor regions (anterior precuneus, mid-cingulate and supplementary motor area), but also in the ventromedial prefrontal cortex, varied depending on whether the self or a friend was being imagined.

Cardiac signals could thus contribute to a body-centered reference frame, to which the brain would refer to in order to tag thoughts as being self-related. We propose that this could be a mechanism for implementing the self, confirming the initial hypothesis that the self is grounded in the neural monitoring of the body.

Résumé

Les théories sur le soi ont postulé que celui-ci serait ancré dans le suivi des signaux viscéraux par le cerveau. Cependant, peu de preuves expérimentales soutiennent ce postulat. L'objectif principal de cette thèse était de traiter directement cette question, en testant expérimentalement si on peut trouver un lien entre le couplage cœur-cerveau et le soi. Nous avons opérationnalisé le concept de soi en définissant deux dimensions du soi : le « Je », expérientiel, et le « Moi », introspectif.

Nous avons montré dans une première étude en magnétoencéphalographie, que le rapport au soi des pensées spontanées est encodé dans l'amplitude des réponses évoquées aux battements cardiaques (*heartbeat-evoked responses*, HERs), dans les régions médiales du réseau du mode par défaut. Plus précisément, les HERs dans le cortex cingulaire postérieur et precuneus ventral encodent le « Je », alors que la dimension « Moi » est associée à des HERs dans le cortex préfrontal ventromédian. Nous avons également montré que ces résultats sont spécifiques de chacune de ces dimensions du soi, ce qui démontre une distinction au niveau biologique entre le « Je » et le « Moi ».

Dans une deuxième étude, nous avons répliqué et étendu ces résultats à l'aide d'enregistrements intracérébraux chez des patients épileptiques. Nous avons montré une covariation entre l'amplitude des HERs et le rapport au soi des pensées, essai par essai. De plus, une analyse par région d'intérêt de l'insula antérieure droite a démontré que les HERs dans cette région sont modulés par la dimension « Je ».

Dans une tâche d'imagination, nous avons trouvé que dans les régions motrices médiales, mais aussi dans le cortex préfrontal ventromédian, l'amplitude des HERs varie en fonction de la personne imaginée, soi-même ou un ami.

Les signaux cardiaques pourraient donc contribuer à l'établissement d'un référentiel centré sur le corps, qui serait utilisé par le cerveau pour attribuer un « label soi » aux pensées. Nous proposons que ceci pourrait constituer un mécanisme pour l'implémentation du soi, confirmant ainsi l'hypothèse initiale selon laquelle le soi est ancré dans le suivi des signaux corporels par le cerveau.

Contents

I. General introduction 15

II. The self 17

A. How to define and study the self 17

B. The neuroscience of the self 18

1.Self-recognition across modalities: towards a unified representation of the self? 18

2. From personality traits to self-reflection 20

3. Building the self in time: autobiographical memory 21

C. The default-network, self-processing and spontaneous thoughts 24

1. Characterization of the default-network 24

2. The overlap between self-related processing and the default-network 26

3. Neural correlates of spontaneous thoughts 27

a) The relevance of spontaneous thoughts 27

b) Intrinsic and extrinsic modes of attention 27

c) How has the self-relatedness of thoughts been probed and what is its relationship with DN activity? 28

D. The "I" and the "Me": from philosophy to cognitive neuroscience 30

1. The "I" and the "Me": what is it? 30

a) Defining the "I" and the "Me" 30

b) The idea of a narrative self 31

c) The "I" and the immunity principle 32

2. Re-interpretation of the neuroscientific findings 32

3. The "I" and the "Me" in cognitive neuroscience 33

III. The self and the living body 36

A. The embodied self 36

B. How malleable is bodily self-consciousness? 38

1. Body ownership and self-location 38

2. First-person perspective 42

3. Agency 46

C. A distinction between experience and introspection relative to the body 48

D. Interactions between bodily processes and higher order self 50

E. From the somatosensory/motor body to the visceral body 51

IV. The visceral body 53

A. Theoretical considerations about the visceral body and the self 53

1. A bodily-centered reference frame for the self 53

2. About visceral signals 54

B. Pathways from the viscera to the brain 54

C. Resting state cortical activity and physiological signals 57

D. Heart and brain 58

1. Stimuli processing and the timing of the cardiac cycle 58

2. Heartbeat-evoked responses 59

a) Origin of the cardiac signal that reaches the brain 60

b) Characterization of the HER waveform 60

c) HER amplitude modulation by different cognitive factors 62

3. Cardiac interoception 64

a) Cardiac interoception measures 64

b) Neural correlates 65

c) Relationship between cardiac interoception and cognition 66

d) What cardiac interoception does and does not tell us 67

E. A major role of the insula? 69

F. Visceral signals and bodily awareness 70

V. Article I: Neural responses to heartbeats in the default-network encode the self in spontaneous thoughts 74

A. Technical remarks on heartbeat-evoked responses 74

1. Confounding artefacts: cardiac-field and pulse artefacts 74

2. Correction and control of artefacts 75

B. Abstract in French 77

C. Article 77

VI. Article II: Is the cardiac monitoring function related to the self in both the default-network and right anterior insula? 90

A. Abstract in French 90

B. Article 90

VII.Article III: Imagining the self is associated with neural responses to heartbeats in medial motor regions and the ventromedial prefrontal cortex 114

A. Abstract in French 114

B. Article 114

VIII. General discussion 137

A. Main results and discussion on the consistency between tasks 137

1. HERs encode the self in spontaneous thoughts 137

2. HERs distinguish self- and other-imagination 138

B. What do these results tell us about the self? 139

1. The “I” and the “Me”: two distinct and graded dimensions of the self in spontaneous thoughts 139

2. What is contrasted when we compare Self and Other? 139

C. Consistency of the results between tasks 140

D. What do these results tell us about spontaneous vs oriented thoughts? 143

E. Proposal of a mechanism for the implementation of the self 143

1. What is this signal? 143

2. Three hypotheses to explain the link between HERs and the self 144

IX. Appendix 148

A. Article I: Phase-amplitude coupling at the organism level: the amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm 148

B. Article II: The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision 158

References 187

I. General introduction

The idea that the self is grounded in the body has been mostly addressed in philosophy. In the neuroscience literature, explicit bodily self-consciousness has been largely explored, but more cognitive or higher-level aspects of the self have been considered mostly independently from the body. Little is known about how the self, taken in its most general definition, could be linked to brain-body interactions. The main goal of this thesis was to study precisely that relationship between brain-body interactions and the self: is there a brain-body mechanism that implements the self?

Our main hypothesis was that signals coming from the body could contribute to the implementation of a body-centered referential defining of the self. This referential would be used by the brain to tag processes as subjective or self-related. Rather than looking for self-specific regions, we propose that we should look for a self-specifying *mechanism* based on ascending bodily signals. To test these hypotheses, we focused on the coupling between the heart and the brain, by measuring brain responses to heartbeats. Are heartbeat-evoked responses related to the self?

Another goal of this thesis was to address the question of the self in a comprehensive way, by specifying the notions of the agentive "I" and the introspective "Me". We hypothesized that these two self-dimensions could underlie most of the more complex forms of self, and that they could both be implemented via neural responses to heartbeats.

In our first experiment, we measured heartbeat-evoked responses during spontaneous thoughts (article I, (Babo-Rebelo et al. 2016a)). We observed that their amplitude differently encoded the "I" and the "Me" in spontaneous thoughts. These MEG results were later replicated and extended, in a study using intracranial recordings from epileptic patients (article II, (Babo-Rebelo et al. 2016b)), aiming at specifying the respective roles of the default-network and insula.

We then questioned to what extent the neural correlates of the self in spontaneous thoughts would resemble those observed during oriented thoughts. We performed another MEG experiment where participants had to imagine themselves or a friend (article III – in prep.). These results show that HERs differ for self- vs other-imagination.

During the course of this thesis, I also had the opportunity to contribute to a review paper, where the idea of a body-centered referential for selfhood is applied to visual perception (appendix, article II - submitted), to an MEG study showing how the gastric activity can constrain the alpha rhythm of the brain (appendix, article I (Richter et al. 2017)) and to ongoing work on single unit responses to heartbeats in intracranial recordings. Additionally, this experimental work led to an opinion review for a French general audience magazine (Cerveau & Psycho, March 2017).

Before presenting the experimental work, we will review the important aspects of the literature that helped define the questions addressed in this thesis.

II. The self

A. How to define and study the self

Although it is very easy to have an intuition about what the self is - our most essential and intrinsic nature, defining it precisely is much harder. Some authors, in particular Metzinger, find the concept so hard to define that they claim the self is no more than a theoretical entity, with no reality and no explanatory function (Metzinger in (Gallagher & Zahavi 2008)).

As a way of apprehending the self, several philosophers and more recently several neuroscientists have proceeded step by step, by decomposing the self into separate dimensions. William James famously described the physical, the social and the spiritual self (James 1890). The physical self not only includes the body, the "innermost part of the material Self", but our clothes, our immediate family, our home… The social self comes from the recognition one gets from acquaintances, making our social selves as numerous as individuals that know us. The spiritual self is the most intimate part of the self, "a man's inner or subjective being, his psychic faculties or dispositions"; it is the source of feelings and thoughts and includes one's personality and values.

Since James, many authors have developed their own partitions of the concept of self. Strawson listed up to 25 forms of self: cognitive, conceptual, contextualized, core, dialogic, ecological, embodied, emergent, empirical, existential, extended, fictional, full-grown, interpersonal, material, narrative, philosophical, physical, private, representational, rock bottom essential, semiotic, social, transparent and verbal (Strawson in (Legrand & Ruby 2009)). Some of these self-dimensions are clearly very closely related (the material and physical selves), but others are less so, for instance the social and embodied selves. If the self has so many dimensions, how can we explain our experience of a unitary self? Gallagher proposes a pattern theory of the self, where all these different dimensions dynamically interact to constitute the self (Gallagher 2013). Thereby, the self is not reducible to any of these features and results from the modulation of the weights attributed to each.

The multifaceted nature of the self is also represented in the diversity of neuroscientific studies linked to it. In the Neurosynth database (Yarkoni et al. 2011),

which assembles the words most frequently employed in cognitive neuroimaging papers, the word "self" is associated with 903 papers (in March 2017). Quite surprisingly, it is almost the same number as for "working memory" (n=901) or "perception" (n=1041), which are more popular domains of cognitive neuroscience. Such a high number for "self" is explained by the diversity of studies which relate to it: emotional processing, theory of mind, autobiographical memory, spontaneous thoughts, personality, own name, metacognition, self-recognition, body, agency, perspective taking… While the theoretical partitions of the self remained quite abstract, the diversity of the neuroscience of the self is expressed in a diversity of cognitive processes. For Gillihan and Farah, this diversity offers the possibility of generating *a posteriori* a full definition of what the self is in reality, encompassing all these different aspects (Gillihan & Farah 2005).

Purpose of this chapter

In the first chapter of this thesis, we will review studies targeting the self as we *know* it: self-recognition (face, body and own name), self-judgment and autobiographical memory. We will discuss the existence of a common neural basis for these different aspects of self (the default-network), and then see how we can look at the self in a more comprehensive way, by specifying the "I" and "Me" dimensions of the self.

B. The neuroscience of the self

1. Self-recognition across modalities: towards a unified representation of the self?

Recognizing one's own face is considered one of the first signs of self-consciousness and appears in children around two years of age (Amsterdam 1972, Rochat 2003). Chimpanzees (Gallup 1970), orangutans (Suarez & Gallup 1981), and even dolphins (Reiss & Marino 2001) and elephants (Plotnik et al. 2006) can presumably pass the mirror test that probes self-face recognition. However, both the validity of this test and the deduction that these animals have self-awareness

continue to be debated (Epstein et al. 1981, Gillihan & Farah 2005, Suddendorf & Butler 2013).

A right hemispheric dominance for self-face recognition has been observed in a number of studies (Keenan et al. 2001, Uddin et al. 2006), but not confirmed in others that found left hemispheric dominance (Turk et al. 2002) or the involvement of both hemispheres (Platek et al. 2006). A recent meta-analysis included 23 articles and confirmed a right hemispheric dominance in self-face recognition (Hu et al. 2016). This meta-analysis further identifies a core set of areas that are more active for self- than for other-face processing, including visual/sensory areas, along with the bilateral inferior frontal gyrus, the right anterior cingulate cortex and the bilateral insula (Figure 1A).

Apart from self-face recognition, the inferior frontal gyrus was shown to be activated by one's own voice (Kaplan et al. 2008) and for one's own moving body parts (Sugiura et al. 2006). Other authors found that the right anterior insula and the right anterior cingulate cortex were implicated in self-face and self-body recognition (Devue et al. 2007). Therefore, these brain regions seem to be implicated in self-recognition, regardless of the modality.

Own name also refers to self-identity and to self-recognition. This particularly salient stimulus attracts attention very efficiently (Wood & Cowan 1995). The effect of one's own name is better studied in electroencephalography odd-ball paradigms, where own name appears rarely, embedded in a stream of other stimuli. The P300 associated with own name is larger than the one associated with other's name, in the right superior temporal sulcus, precuneus and medial prefrontal cortex (Perrin et al. 2005). In fMRI, midline cortical structures were shown to be more active in the detection of own name than other's name (Kampe et al. 2003), even in six-month-old infants (Imafuku et al. 2014). Considered as a proxy of self-consciousness, this oddball paradigm has been used to assess the level of consciousness in minimally conscious, vegetative state and locked-in patients (Perrin et al. 2006). Interestingly, an electrophysiological correlate of own name has also been found in chimpanzees (Ueno et al. 2010).

Conclusion

From the results of visual recognition and name detection, it is not yet clear whether we can really define a cross-modal network of self-recognition. Some regions seem to be responsive to different kinds of self-related stimuli, such as the inferior frontal gyrus, the medial prefrontal cortex / anterior cingulate.

The ability to recognize oneself is considered a fundamental building block of self-awareness, which can be based on different modalities (self-face, body, voice, name…). Self-recognition would be the first step toward the ability to introspect, i.e. to become the object of one's own attention (Gallup et al. 2014). Let us review now the main neuroscientific findings concerning self-reflection.

2. *From personality traits to self-reflection*

Self-reflection refers to the capacity of thinking about oneself, of judging oneself. One of the paradigms involving self-reflection is the judgment of personality traits in relation to self or other.

Both self- and other-reflection activate midline cortical structures, in particular the medial prefrontal cortex, the left temporoparietal junction, the posterior cingulate / precuneus, the left middle temporal gyrus and the superior temporal sulcus (Denny et al. 2012, van der Meer et al. 2010). As compared to other-judgments, self-judgments elicit an increase in activity especially in midline structures such as the anterior cingulate cortex / medial frontal gyrus and the precuneus (Denny et al. 2012, Hu et al. 2016, van der Meer et al. 2010) (Figure 1B).

Differences between self and other depend on the familiarity of the other (Figure 1D, E). When compared to reflection about a distant other (the former US president G. W. Bush or the Danish queen), self-reflection activates these regions more (Kelley et al. 2002, Kjaer et al. 2002, Schmitz et al. 2004). This difference is less clear for the comparison between self and a close other (Murray et al. 2012). One possible explanation is that processing information about a close other engages self-referential processing as well (Jenkins et al. 2008).

Interestingly, the brain regions activated for self-reflection about personality traits are also activated for reflection about one's current mental states or physical

attributes (Jenkins & Mitchell 2011), as well as reflection about one's own feelings (Ochsner et al. 2004).

The ventromedial prefrontal cortex in particular appears to be more associated with self- than other-judgement (van der Meer et al. 2010) whereas the dorsal medial prefrontal cortex is associated with other-judgment, suggesting a self-other gradient in the medial prefrontal cortex (Denny et al. 2012). For D'Argembeau, the ventromedial prefrontal cortex would evaluate any kind of self-related information in order to assign personal value (D'Argembeau 2013).

Activation of the posterior cingulate cortex has been interpreted as a sign of autobiographical memory retrieval (van der Meer et al. 2010), which might occur during trait judgment tasks (Araujo et al. 2013) (Klein 2004).

Conclusion

Reflecting about oneself is associated with activity in midline brain regions, regardless of modality, suggesting the existence of a common underlying mechanism of self-reflection. While the ventromedial prefrontal part would be associated with the assignment of subjective value, the posterior part would be associated with retrieval of autobiographical memories. Indeed, as personality is built up over a lifetime, autobiographical memory constitutes an important building block of the self.

3. Building the self in time: autobiographical memory

Autobiographical memory can be defined as "the memory systems that encode, consolidate and retrieve personal events and facts" (Fossati 2013). Autobiographical memory defines the self across time, thereby constituting one's identity. Interestingly, some brain regions are responsive to both present and past selves (past and present self-face (Apps et al. 2012), reflection about the past and present selves (D'Argembeau et al. 2008)), which supports the idea of the continuity of the self in time.

Retrieving memories engages a variety of cognitive functions, such as emotional processing, executive control, visuospatial processing, working memory, attention or self-processing. This may explain the diversity of brain regions

associated with the retrieval of autobiographical memories. Indeed, a meta-analysis on 24 studies on autobiographical memory has identified a core network of regions including midline cortical structures (medial prefrontal, retrosplenial/posterior cingulate and medial temporal cortex – hippocampus, parahippocampus, perirhinal and entorhinal cortices), as well as lateral prefrontal and lateral temporal regions, temporoparietal junction and the cerebellum (Svoboda et al. 2006). Midline structures in particular have a causal role in memory retrieval, since patients with lesions in these regions demonstrate impairments in retrieving memories from their lives (Philippi et al. 2015).

Midline cortical structures are particularly interesting because they process autobiographical memory retrieval in a gradient of increasing abstraction, from posterior to anterior regions (Martinelli et al. 2013) (Figure 1C). Episodic autobiographical memory (retrieval of memories of specific events) is associated with precuneus/posterior cingulate cortex; semantic autobiographical memory (retrieval of general personal events and personal information) with medial prefrontal and posterior cingulate cortices; and the conceptual self (personality-trait judgment) with ventromedial and dorsomedial prefrontal and anterior cingulate cortices (Martinelli et al. 2013).

Cortical midline regions would mediate self-referential processing (Summerfield et al. 2009), whereas a parieto-temporal subsystem would be more associated with memory retrieval (Kim 2012).

Conclusion

Autobiographical memory supports the continuity of the self along time. Retrieving a past episode from memory engages many different cognitive functions, but the self-referential processing appears to be ensured by the activity of midline cortical structures.

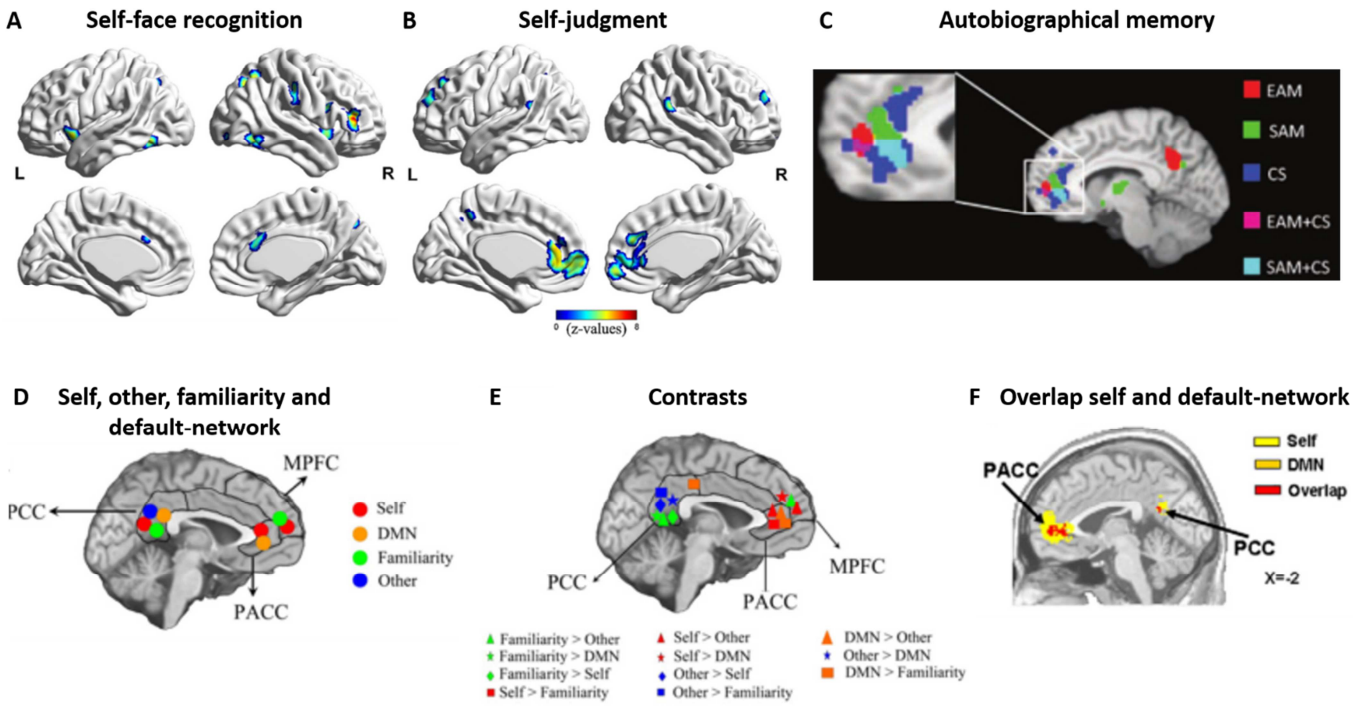


Figure 1

Brain regions involved in different self-related processes.

(A) Brain regions involved in self-face recognition (self-face > other–face), according to a meta-analysis of 23 articles (Hu et al. 2016)**.**

(B) Brain regions involved in self-judgment tasks (mainly personality trait judgment, self-judgment > other-judgment), according to a meta-analysis of 37 articles (Hu et al. 2016)**.**

(C) Brain regions involved in autobiographical memory retrieval (EAM episodic autobiographical memory, SAM semantic autobiographical memory, CS conceptual self), according to a meta-analysis of 38 articles (Martinelli et al. 2013).

(D) Summary of the results of a meta-analysis on the default-network (DMN, 24 papers), self- (57 articles), familiarity- (23 articles) and other-processing (23 articles) (Qin & Northoff 2011). PCC posterior cingulate cortex, MPFC medial prefrontal cortex, PACC perigenual anterior cingulate cortex.

(E) Results of the comparison between the different conditions shown in (D) (Qin & Northoff 2011).

(F) Overlap between region involved in self-processing and the default-network (Qin & Northoff 2011).

Conclusion of B. and our proposal

We have reviewed some of the literature on the neural bases of self-recognition, self-judgment and autobiographical memory retrieval. These processes constitute different features of self-consciousness, which are profoundly intertwined. For instance, reflecting about oneself requires self-recognition and can be based on personal memories from the past. From the results cited above and from Figure 1, we can see that cortical midline structures are involved in most of these self-related processes, but these are not the only structures involved. Moreover, they are – to some extent – responsive to other-related stimuli as well.

Should we then question the unity of the self, based on the fact we do not find regions specifically dedicated to self-related processing? Instead of thinking in terms of *overlapping and specific brain regions*, we propose that we should look instead for a *common mechanism* implementing the self in different brain regions. To develop this hypothesis, we need first to better characterize these midline cortical regions (part C) and then try to operationalize the self in a comprehensive way (part D), underlying the different kinds of self-related processes we discussed in part B.

C. The default-network, self-processing and spontaneous thoughts

1. Characterization of the default-network

We have seen that midline cortical structures play a major role in the self-related processes described earlier. These regions are part of the default-network (DN[1]) (Figure 2). The overlap between self-processing and DN has been noted in many studies (Goldberg et al. 2006, Gusnard et al. 2001, Kim 2012, Schneider et al. 2008). We will first characterize this resting state network and, in the next section, examine how it relates to self-processing.

[1] We here use "default-network" instead of using the more common name of "default-mode network", following (Andrews-Hanna et al. 2014): "the latter refers to passive states, which may obscure the adaptive functions of this network. The former is meant to emphasize its role as a large-scale brain system whose functions may extend beyond the resting state".

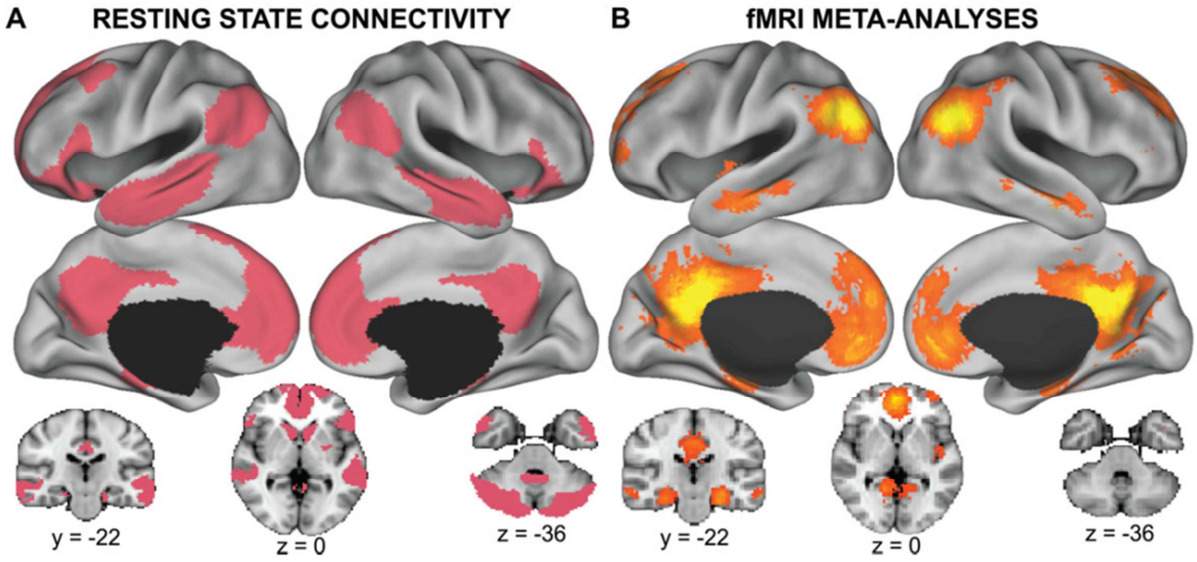


Figure 2

The default-network.

(A) The default-network as revealed by resting-state functional connectivity MRI of the cortex, striatum, and cerebellum.

(B) The default-network revealed by a meta-analysis of functional neuroimaging data using NeuroSynth software. Shown are false discovery rate-corrected reverse inference statistical maps (P term|activation) for meta-analyses corresponding to default mode, default network, or default mode network.

(figure and legend from (Andrews-Hanna et al. 2014)**)**

The first observations of the DN date from the 50s and 70s, when high levels of metabolism and cerebral blood flow were observed during the resting state (Buckner et al. 2008, Ingvar & Schwartz 1974, Sokoloff et al. 1955). This network was then defined as a set of brain regions which are more active during rest than during task performance, i.e. regions showing task-induced deactivations (Andrews-Hanna et al. 2014, Buckner et al. 2008). The main regions are the medial prefrontal cortex, the posterior cingulate cortex / ventral precuneus / retrosplenial cortex and the inferior parietal lobule. The hippocampus and the lateral temporal cortex are also part of the DN but are less prominent. The DN is now identified according to the patterns of temporal correlations (functional connectivity) between its different regions during rest (Figure 2A) (Fox et al. 2005, Yeo et al. 2011). Although the electrophysiological characterization of the DN is still very preliminary, it shows correlation maps on the source space that are coherent with fMRI results (de Pasquale et al. 2010).

The posterior midline component of the DN (posterior cingulate cortex and ventral precuneus – the dorsal precuneus is not part of the DN (Buckner et al. 2008)) is considered to play a fundamental role in the dynamics of spontaneous brain fluctuations. Indeed, this node is densely connected to many brain regions (Bzdok et al. 2015, Margulies et al. 2009) and is considered a hub (Leech et al. 2012), i.e. a structure that integrates information across different functional networks. It is the only DN node that displays interactions with all the other nodes of the network (Fransson & Marrelec 2008).

2. The overlap between self-related processing and the default-network

The anterior and posterior cingulate cortices belonging to the DN have been associated with several self-related tasks (Qin & Northoff 2011, Spreng et al. 2009) (Figure 1F). Meta-analyses have shown that in these tasks (trait adjective judgment, face or body recognition, name perception, autobiographical memory…), the contrast self vs other yields larger activations in these regions for the condition self (Figure 1E), but also activates these regions relative to a baseline in the condition other (Qin & Northoff 2011).

A direct comparison of activations during a self-reference task (trait adjective judgment) and resting state showed that default network regions (in particular the medial prefrontal cortex and the posterior cingulate) were indeed active during both (D'Argembeau et al. 2005, Davey et al. 2016, Whitfield-Gabrieli et al. 2011). The cerebral metabolism in the ventromedial prefrontal cortex, in particular, correlates with the amount of self-referential thoughts (D'Argembeau et al. 2005).

A possible reason for this overlap between DN regions and self-processing is that spontaneous thoughts during the resting state are very often self-related. We will now review evidence showing the relationship between the self, mind wandering and the DN.

3. *Neural correlates of spontaneous thoughts*

a) The relevance of spontaneous thoughts

DN activity has been associated with the state of mind wandering (Fox et al. 2015). The initial definition of the DN as being "task-negative" led to the wrong idea that the DN did not have much of a cognitive function. Even though the DN is deactivated during externally-oriented tasks, it is active during rest when self-generated cognition and mind wandering take place (Andrews-Hanna et al. 2010a, Christoff et al. 2009, Mason et al. 2007).

Mind wandering is actually one of the most prominent mental activities in our daily lives. We all know that spontaneous thoughts can interrupt the concentration on a task we are trying to perform. In a large survey using an iPhone app that probed users at a random times during the day (Killingsworth & Gilbert 2010), Killingsworth and Gilbert showed that 46.9% of the probes caught people in a mind wandering state.

These spontaneous thoughts often concern past events that are remembered, or future events that are imagined, with a bias towards future (prospective bias (Smallwood & Schooler 2015)) and pleasant events (Killingsworth & Gilbert 2010). Spontaneous thoughts can adopt several forms, they can involve mental imagery, inner language, bodily awareness, music (Delamillieure et al. 2010). Inner thoughts would have multiple adaptive advantages such as allowing us to better prepare future events by simulating them, solving one's concerns, navigating our social world and developing a form of self-identity that connects our past, present and future selves (Andrews-Hanna et al. 2014).

b) Intrinsic and extrinsic modes of attention

Some studies have explored the relationship between the content of thoughts at rest and brain activity, by distinguishing between internally- and externally-directed awareness. Internally-directed thoughts, as defined in (Vanhaudenhuyse et al. 2011), relate to inner speech, experiment-related or autobiographical thoughts, as opposed to externally-directed thoughts, which correspond to thoughts elicited by auditory, somesthetic, olfactory or visual stimuli.

Internal awareness was associated with activity in midline DN structures (anterior cingulate, posterior cingulate, parahippocampal cortices), whereas external awareness was associated with more lateral regions (inferior parietal lobule, dorsolateral prefrontal cortices) (Vanhaudenhuyse et al. 2011) and sensory regions (Golland et al. 2007, 2008; Tian et al. 2007). The activity of these systems is anti-correlated and could reflect behavioral competition between being on- and off-task; the former being associated with extrinsically-driven attention, the latter with mind wandering (Christoff et al. 2016, Schooler et al. 2011). Some authors have even considered that the "brain loses its self" during the performance of a demanding perceptual task, because internally-oriented processes would be suppressed (Goldberg et al. 2006).

Attention that is decoupled from the environment can be related to the self if spontaneous thoughts involve autobiographical memory or introspection for instance. However, this internal/external distinction does not correspond neatly to a self/non-self distinction, since for instance an unpleasant feeling provoked by a sound would be considered externally-directed while it is highly self-relevant. Conversely, thinking about the political state of the world can be a thought independent from the direct surrounding environment and still not self-related. How can we directly assess the self-relatedness of thoughts?

c) How has the self-relatedness of thoughts been probed and what is its relationship with DN activity?

Andrews-Hanna and colleagues explored the self-relevance and time orientation of thoughts in order to define specific contributions of different DN nodes (Andrews-Hanna et al. 2010b, Christoff et al. 2016). A core network composed of the posterior cingulate and the anterior medial prefrontal cortex encodes the self-relevance of thoughts, associated with its affective value, regardless of time orientation. Imagining a scene based on memories is in turn associated with activity in a medial temporal subsystem (retrosplenial cortex, ventromedial prefrontal cortex, posterior inferior parietal lobule, parahippocampal cortex and hippocampal formation). Thinking about one's present mental state activates the dorsal medial prefrontal cortex subsystem (dorsal medial prefrontal cortex, temporoparietal junction, lateral temporal cortex and temporal pole).

Like in (Tusche et al. 2014), the characterization of the self-relatedness of thoughts in (Andrews-Hanna et al. 2010b) is quite vague (just self-related or other-related). In turn, in (Andrews-Hanna et al. 2013), the self-relatedness of thoughts was decomposed into so many categories, that some actually overlap. For instance, self-relevance was highly correlated with centrality and subjective value possibly because all involve an introspective point of view towards oneself.

Our proposal

In our view, the most parsimonious and comprehensive way to describe the self-relatedness of thoughts is to define the "position" of the self in the thought. Was I thinking that I was performing some action? Was I introspecting about myself? This corresponds to the distinction between the "I" (self as the subject, the agent) and the "Me" (self as the object of introspection). In spontaneous thoughts, this distinction is rather equivalent to the grammatical distinction between the subject and the object of the sentence. The "I" is engaged when one is adopting the first-person perspective, when one is the agent, the one doing something in the thought. For instance, the "I" is engaged in a thought like "I will go to the supermarket", but not in "He is going to the supermarket". On the other hand, the "Me" refers to thoughts where one is thinking about oneself, where one is introspecting, thinking about his/her feelings, bodily state, as in "I am tired". The self can thus be expressed in two different – but not exclusive – ways, so that any spontaneous thought can be reduced to these two dimensions.

Theoretical support for the relevance of this distinction as well as how it can be applied in experimental cognitive neuroscience will be reviewed in the following section.

D. The "I" and the "Me": from philosophy to cognitive neuroscience

1. *The "I" and the "Me": what is it?*

Despite disagreements concerning the concept of self, there is a general consensus in the history of philosophy, especially among phenomenologists, that a distinction can be made between the self-as-subject, the "I", and the self-as-object, the "Me", and in considering that the self is not reducible to the self-as-object dimension (Christoff et al. 2011, Legrand & Ruby 2009).

a) Defining the "I" and the "Me"

During self-reflection, the self is the *object* of introspection; consciousness is directed towards oneself. This "Me", or self-as-object, is explicit, linguistic and can be narrative (Table 1). In contrast to this reflective self-awareness, Sartre defines a pre-reflective self-awareness (Sartre in (Zahavi 2005) p21). In every experience, there is a self who is the *subject* of experience (i.e. the "mineness" of experience). Similarly, William James opposes the self as the "Me", "matter" of thoughts, with the "I", the self as the "thinker" (James 1890).

To take the words of Husserl (Husserl *in* (Zahavi 2005), p117), the *intentional quality* of the experience can vary (the type of experience: perceiving, remembering, doubting…) as well as its *intentional matter* (the object of experience: an experience of a cat, of a tree, memory A or memory B…), but the first-person experiencing subject remains invariant. It is the only aspect that remains constant throughout all kinds of experiences (Legrand 2007, Zahavi 2005). This self is pre-reflective, since experiences are intrinsically and implicitly experienced from the first-person perspective (Legrand 2007) (Table 1). Legrand defines the self-as-subject as being "neither an external object (for example, it is not my body that I can observe in the mirror) nor an internal object: when I am conscious of myself as the subject of an experience, I am not scrutinizing an internal self looking at the external world. I am simply looking outside at the external world, and within this single act of consciousness I pre-reflectively experience myself-as-subject". Furthermore, still according to Legrand and in agreement with most phenomenologists, "pre-reflective

self-consciousness is the necessary ground upon which other forms of self-consciousness are anchored", including the self-as-object (Legrand 2007).

b) The idea of a narrative self

Related to the "I" / "Me" distinction, other authors have proposed distinguishing the minimal from the narrative self. The narrative self of Gallagher is "a more or less coherent self that is constituted of a past and a future in the various stories that we and others tell about ourselves". In contrast, the minimal self is "phenomenologically, [...] a consciousness of oneself as an immediate subject of experience" (Gallagher 2000). Here, time seems to be the crucial difference between the two: the minimal self being the present and immediate self (the synchronic self), while the narrative self is the self of the past or of the future (the diachronic self), that has to be mentally re/pre-constructed. For Dennett, the self is nothing but a narrative construal, a product of imagination, a fiction, a center of gravity reuniting all the stories about the self, but with no true meaning (Dennett 1991). Zahavi argues that still the narrative self requires a first-person perspective narrator, that enables the self/non self distinction, the self-attribution of actions and agency and the ability to use the first-person pronoun (Zahavi 2005). Zahavi proposes saving the word "self" for the pre-reflective self, and "person" for the narrative self, which contains the history, the personality, the identity of the subject acquired during their lifetime. The *person* can be reflected upon (self-as-object) whereas the *self* is implicitly present in our stream of consciousness (self-as-subject).

"I"	"Me"
Self-as-subject	Self-as-object
Pre-reflective self	Reflective self
Minimal self (present)	Narrative self (extended in time) when introspected
Implicit	Explicit
Experiential	Linguistic
First-person perspective experience	Introspection, self-reflection
Grounds the "Me"	Grounded in the "I"
Immune to error through misidentification	Not immune to error through misidentification

Table 1

Concepts underlying the "I" and the "Me" and their different properties.

c) The "I" and the immunity principle

The first-person pronoun "I" has an important feature that the "Me" does not possess: the immunity to error through misidentification (IEM) (Table 1), according to Wittgenstein and Shoemaker (Gallagher 2000, 2012). This means that when we use the first-person pronoun "I" we cannot be mistaken about whom it is referring to, since the "I" is intrinsic to the experience we are preparing to talk about. If I am experiencing a toothache, it does not make sense to ask "Someone has a toothache, is it I?". I cannot be mistaken about the fact I am the one having the toothache because the "I" is built into the experience itself. In contrast, if I see a sunburned arm in a mirror, I may misattribute it to myself, when in fact it is someone else's arm. The fact that the "I" is immune to this kind of error shows how basic and essential this form of self is, in contrast to the "Me".

2. Re-interpretation of the neuroscientific findings

Most of the paradigms we reviewed in chapter B "The neuroscience of the self" are based on the self-attribution of mental or physical features: attribution of personality traits, faces or names. As participants have to think about themselves the self is the *object* of attribution ("Me"). For Legrand and Ruby this process is not *self-specific* (Legrand & Ruby 2009). Self-specificity is defined by two criteria:

- Exclusivity: a self-specific component characterizes the self but not the non-self,
- Non contingency: changing or losing this component leads to a change or loss of the distinction between self and non-self.

Understanding the neural bases of the self would require understanding *self-specific* components of the self, and not just the distinction between the processing of self-related and other-related contents. Personality traits are not self-specific because they can characterize the other, so they do not meet the criterion of exclusivity. Traits and self-face recognition are also not self-specific because they are contingent: one's personality or one's own face can change, but this does not imply a loss of the self/non-self distinction. Rather than being in the *content* of the process, for Legrand and Ruby, self-specificity is in the *perspective* of the process. The perspective is what links the perceiving *subject*, the "I", to the perceived object. Experiences are self-specific, because the experience I have, from my perspective, is systematically

different from the experience someone else has, thereby meeting the criterion of exclusivity. Moreover, perspective is non contingent, we cannot lose it: one can take a third-person perspective, but the origin is always that of first-person.

Legrand and Ruby (Legrand & Ruby 2009) therefore reinterpret the neuroscientific results in terms of evaluation and inference processes. These paradigms would involve a process of *evaluation*: integration of the stimulus, memory recall, comparison of memories and stimulus, and finally integration of all elements to form a conclusion. These evaluative steps are engaged irrespective of the subject of the task (self or other), which could explain why regions associated with the self are also, to some extent, responsive to other. In their view, differences between self- and other-processing would stem from differences in the need for inferences (which would correspond to medial prefrontal activations) and in memory recall (which would correspond to posterior cingulate cortex activations) between self and other conditions.

Therefore, the distinction between the "I" and the "Me" has important consequences in the interpretation of neuroscientific results.

3. The "I" and the "Me" in cognitive neuroscience

As said earlier, experiments like the personality trait judgment or the self-face recognition, and to some extent even autobiographical memory recollection target the self as a self-related content, i.e. the "Me". The "I" remains to be addressed as such, but since it is the ground for the self-as-object, it is somehow present in these experiments too.

The term "self-consciousness" might have given rise to this bias in the neuroscientific literature towards the self-as-object. "Self-consciousness" is misleading in the sense that it sounds equivalent to "consciousness *of* self", which refers to self-reflection (to the "Me") and leaves aside the implicit experiencing subject (the "I"). Sartre even proposed writing "consciousness (of) self" to stress that "of" is a necessary (but misleading) grammatical formulation (Sartre *in* (Zahavi 2005)).

While we can easily conceive of what the experimental contrast for the self-as-object could be (a content related to someone else, like someone else's face), it is

harder to find an experimental contrast for the self-as-subject. Indeed, it is not clear in the phenomenological concept of self-as-subject what the absence of self-as-subject could be. Not only is it immune to errors, but it also seems to underlie every event of our mental lives: it implicitly underlies every kind of perception, thought, first-person perspective and even the adopting of a third-person perspective (Legrand 2007).

As we exposed earlier, every experience is intrinsically and necessarily linked to an experiencing subject. This might imply that either there is nothing to understand about the experiencing subject in biological terms (i.e. understanding the neural mechanisms of visual perception, for instance, is also understanding the neural mechanisms of the perceiving subject) or that there is still a biological mechanism intrinsically associated with perception that implements the perceiving subject.

In visual perception experiments, subjects are typically asked to report what they see. Participants therefore experience some visual perception (they are the experiencing subject) but they also reflect on their own experience. Frässle and colleagues have shown that pure perception and report are indeed distinguishable. During a binocular rivalry task, frontal activations were associated with the online report of the percept, while passively experiencing the alternating percepts activated only occipital and parietal regions (Frässle et al. 2014). This shows that experience and introspection about experience can be dissociated in visual perception, which fits with the idea of an experiencing “I” being different from an introspective “Me”.

Summary, proposals and conclusion of part II

In this chapter we reviewed neuroimaging studies concerning self-recognition, self-judgment and autobiographical memory. Midline cortical regions are consistently involved in self-related processing, along with other regions which are less consistent across studies. We proposed that, instead of characterizing the self through regions where different self-processes overlap, we should look instead for a common brain mechanism underlying the self.

This mechanism could be at play in midline cortical regions, which are part of the default-network. This network is highly active during the resting state and relates to the content of spontaneous thoughts. Because spontaneous thoughts are often self-related, this could explain why we find these regions in self-related tasks. The self can be expressed in different ways during spontaneous thoughts, which makes it difficult to apprehend. We proposed that we could distinguish the "I" and the "Me", i.e. the agentive self and the introspective self. This distinction is actually based on a vast philosophical literature, which has important consequences in the interpretation of the results from neuroimaging experiments, which usually tend to focus on the "Me" dimension. Therefore, our goal was to try to find a basic mechanism that would define the self at the brain level, and that would characterize the "I" and the "Me".

The self is firstly defined by the body and bodily signals are constantly monitored and integrated in the brain. Our hypothesis was therefore that anchoring the self to the body could be a mechanism by which the self is defined at the brain level.

We will now see how brain-body interactions contribute to the self. We will first talk about the living, somatosensory/motor body (part III), and then about the visceral body (part IV).

III. The self and the living body

A. The embodied self

"Our entire feeling of spiritual activity [...] is really a feeling of bodily activities" (James 1890). William James adds here an important element to the concept of self. For him, the "spiritual self" or any element of our mental lives has a bodily basis. Any perception is accompanied by the adjustment of the corresponding sensory organ (of the eyeballs for vision, for instance); mental effort is accompanied by movements of the brows or eyelids, contractions of jaw-muscles etc. In his hierarchy of the self, "the bodily self [is] at the bottom, the spiritual self at the top, and the extracorporeal material selves and the various social selves between". Every episode of our lives is accompanied by particular bodily feelings that remain associated with the memories we retain from those episodes. These bodily feelings give the memories the "warmth" necessary for memories to be felt as our own and a sense of continuity of the self along time, as being one and the same.

The famous brain-in-the-vat thought experiment has challenged the idea of the necessity of a body. Imagine an isolated brain, in a vat containing all the chemicals necessary for its normal functioning and connected to various electrodes that inform it about the world. Is the brain sufficient for experience and cognition to occur? If yes, then the body is unnecessary. The answer of Gallagher and Zahavi is that the brain-in-a-vat is still dependent upon all the resources usually provided by the body, which are here provided by electrodes (Gallagher & Zahavi 2008). For Damasio, the absence of body-brain loops compromises the emergence of a normal mind. Bodily inputs not only contribute to the normal functioning of the brain, but importantly they are "a *content* that is part and parcel of the workings of the normal mind" (Damasio 1994). Furthermore, it is an empirical fact that cognition *is* embodied, our body allows and shapes our perception and actions. For instance, the shape of the ears explains different auditory capacities among primates (Coleman & Ross 2004). If we think in evolutionary terms, the primary function of the brain is to maintain the body, our mental life comes on top of that (Damasio 1994).

According to the phenomenological tradition, especially Merleau-Ponty, Husserl and Sartre, more than being an object of the world, the body is the "principle of experience", it allows our experiences (Gallagher & Zahavi 2008). In this sense, the

body is “pour-soi” (“being-for-itself”), what I and only I can experience from inside, the lived body; and not just “pour-autrui” (“being-for-others”), its properties and shape that can characterize any body (Sartre 1943). In this sense, face-recognition as we saw in part II.B.1 relates to the “pour-autrui” body. What we want to discuss now is the lived body, the “pour-soi” body.

More recently, Damásio has developed a theory of the self in neuroscientific terms, in line with William James and the phenomenologists, which places bodily processes at the bottom of the hierarchy of the self. The first level of self is the *proto-self*, which represents the unconscious monitoring of the moment by moment state of the body. Anytime an object (a face, a melody…) interacts with the organism, it modifies the organism’s state thereby modifying the proto-self. These new maps representing the interaction between the organism and the object can become conscious and generate the *core-self*. Because a multiplicity of objects is constantly interacting with the organism, the core-self is constantly generated and continuous in time. The third and more elaborated level is the *autobiographical self*, which is the collection of experiences of the core-self. The autobiographical self is extended in time and places the subject at a certain point in their personal history, with a certain past that constitutes their identity, and a perspective of the future. In this hierarchical model of the self, each level depends on the lower level. For Damásio, the body ensures the stability of the self, i.e. this feeling that we stay the same person throughout our lifetime. Because the range of internal bodily states compatible with life is actually limited, bodily representations are stable. Further, while the environment changes continuously, bodily representations remain relatively constant, thereby ensuring the stability of the self.

Blanke and Metzinger also ground self-consciousness in the body (Blanke & Metzinger 2009). For them, the minimal phenomenal selfhood qualifies the phenomenal experience of being a self, and is fundamentally based on a body-centered reference frame, i.e. a bodily representation from which a “weak first-person perspective” emerges. A “strong first-person perspective” occurs when attention is focused on a certain object, which can be the body itself. Here, attentional orienting towards the self is grounded on a basic representation of the body that constitutes the most basic form of self.

Conclusion

Bodily self-consciousness, or the consciousness of the body, is the first step for the link between the self and the body. We will now talk about the lived body, the body as directly experienced from within, the body that shapes our experience of the world. How are we conscious of our body? How malleable is the consciousness we have from our body? How does it influence more global aspects of the self?

We will look at different elements of bodily self-consciousness – from consciousness of the body to interactions between the body and the environment. We will start with self-location and body ownership, two ways of being conscious of one's own body. We will then talk about the first-person perspective, the perspective of the environment that is centered in our body. Finally, we will discuss agency, when the body is owned and controlled so as to perform specific actions having an outcome in the outside world.

B. How malleable is bodily self-consciousness?

1. Body ownership and self-location

Body ownership and self-location are two basic components of bodily self-consciousness. Body ownership refers to the feeling that this body or this body part is mine. Self-location corresponds to the feeling that my body is located in a specific point in space.

Both body ownership and body location can be altered in patients with particular neurological conditions (Blanke 2012). Some patients with somatoparaphrenia do not feel ownership over one specific body part, which they think belongs to someone else. Conversely, some other patients experience hands of other people as belonging to themselves. Patients with autoscopic phenomena (Blanke & Metzinger 2009) report seeing a second own body in extracorporeal space. These full-body illusions induce global changes in bodily self-consciousness.

Following these observations in patients, experimental visuo-tactile illusions have been developed. When viewing a rubber hand being stroked at the same time as one's own hand is stroked, one can experience illusory self-attribution of the rubber

hand and a spatial displacement of one's own hand toward the rubber hand (Botvinick & Cohen 1998, Ehrsson et al. 2004, Tsakiris & Haggard 2005).

This rubber hand illusion can be induced even for a very long arm or for an arm with a different color, but not if the rubber hand is outside of the peripersonal space, if its orientation is not anatomically plausible or if it does not look like a hand (for a review (Blanke et al. 2015)). This shows some plasticity in bodily self-consciousness, as long as the global bodily shape is preserved. This plasticity does not mean that an extra hand is added to our body representation, but rather that the real hand is replaced by the rubber hand – is embodied (Longo et al. 2008). The feeling of ownership of the rubber hand can be so strong that if the rubber hand is threatened participants have a feeling of anxiety as if their own hand was threatened (Ehrsson et al. 2007).

This experiment has also been applied to the face (enfacement illusion) (Sforza et al. 2010, Tsakiris 2008). Participants watch a morphed face being touched in synchrony with strokes applied to their own face. This manipulation induces a bias in a subsequent self-recognition task, where the other person's face is included to a greater extent in the representation of one's own face. It is also possible to induce full-body illusions by stroking the participant's back (Lenggenhager et al. 2007), chest (Ehrsson 2007) or both (Lenggenhager et al. 2009) in synchrony with the strokes applied to a virtual body located in front of the participant. Participants may then experience self-identification with the virtual body (Lenggenhager et al. 2007, 2009) and a drift in self-location (Ehrsson 2007, Lenggenhager et al. 2007), to the extent that threats to the virtual body evoke large skin conductance responses in the participant (Guterstam et al. 2015).

Recently, the rubber hand illusion was adapted to mice, in the form of the rubber *tail* illusion (Wada et al. 2016). When their own tail is stroked synchronously with strokes applied to a rubber tail they can see, mice react to a threatening stimulus directed to the rubber tail as if it was their own tail. This result shows not only that mice may have a form of bodily self-consciousness but also that bodily self-consciousness is truly and intrinsically malleable. This new paradigm opens up important new avenues for research, in particular to understand the electrophysiological correlates of these illusions.

These illusions work when visual and tactile stimuli are synchronous and induce proprioceptive/vestibular changes, suggesting that they are supported by multisensory integration (Blanke 2012, Blanke et al. 2015, Tsakiris 2010). Multisensory integration would be performed at the level of bimodal or trimodal neurons, which are responsive to stimuli in the peri-hand or peripersonal space (Makin et al. 2008). Visuo-tactile stimulations would alter the receptive fields of multisensory neurons, in order to include the virtual body or the fake body part (Blanke et al. 2015). Proprioceptive cues are integrated as well, ensuring that the global body posture is preserved. These multimodal neurons are thought to be part of a multisensory network encompassing the posterior parietal and premotor cortices (Blanke 2012, Blanke et al. 2015, Ehrsson et al. 2004, 2005) (Figure 3). This change in multisensory receptive fields would result in changes of ownership for body parts (rubber hand) and self-identification (full-body illusion) (Blanke et al. 2015).

By integrating multisensory information across body parts (Petkova et al. 2011), the premotor cortex would be responsible for full-body ownership (Guterstam et al. 2015). Self-location would in turn be associated with the hippocampus, posterior cingulate, retrosplenial and intraparietal cortices (Guterstam et al. 2015), as well as the temporoparietal junction (Ionta et al. 2011a). The posterior cingulate cortex would be responsible for the integration of body ownership and self-location, leading to a complete sense of bodily self-consciousness (Guterstam et al. 2015).

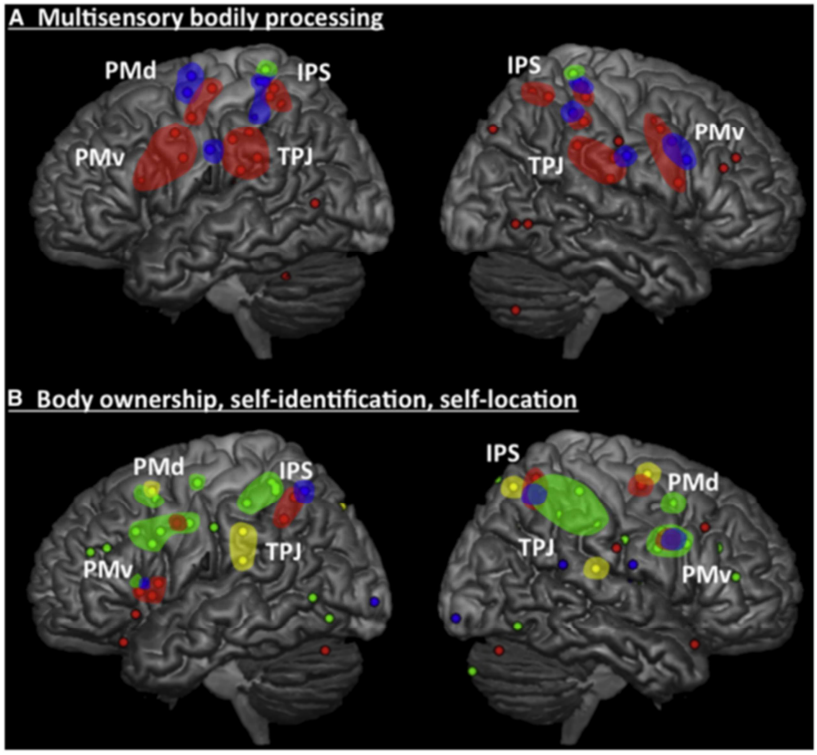


Figure 3

(A) Brain areas selectively responding to multisensory inputs within the peripersonal space around the hand (red), face (blue) or trunk (green).

(B) Brain areas active during manipulations of bodily self-consciousness, underlying ownership for the hand (red) or face (blue), self-identification (green), or self-location (yellow). Each dot represents an activation site as identified by the studies reviewed in (Blanke et al. 2015). The colored shadows highlight clusters of activations.

IPS intraparietal sulcus; TPJ temporoparietal junction; PMd dorsal premotor cortex; PMv ventral premotor cortex.

(figure and legend from (Blanke et al. 2015)).

Conclusion

Pathological cases and bodily illusions show how malleable our bodily self-consciousness is. The rubber hand illusion induces changes only in ownership for one body part and depends on the individual's peripersonal space, while stimulation of the trunk induces a global and unitary illusion that alters more and more radical aspects of bodily self-consciousness. The enfacement illusion in turn impacts the representation one has of oneself. However, these illusions only work with objects that resemble the shape of the body and that respect body posture.

Multisensory integration is the mechanism through which body ownership and self-location are implemented. Tactile and visual stimuli are integrated together with proprioceptive signals into a common reference frame – the subject's body. This system presents the advantage of allowing the embodiment of artificial limbs through prolonged multisensory stimulation, but also leads to global changes of bodily self-consciousness in the case of trunk stimulation.

2. First-person perspective

First-person perspective is a fundamental building block of bodily self-consciousness and is defined as the perspective from where I perceive the world. A first-person perspective coincides with self-location, except in rare cases (De Ridder et al. 2007), and can be altered in the kind of full-body illusions described above. Interestingly however, self-location (Ionta et al. 2011b, Pfeiffer et al. 2013) and body ownership (Fotopoulou et al. 2011, Petkova et al. 2011) appear to be associated with the first-person rather than the third-person perspective. Therefore, the first-person perspective, self-location and body ownership are very intertwined and usually congruent, but can be dissociated in some cases.

Contrarily to out-of-body experiences and full-body illusions which can alter general bodily self-consciousness, a voluntary change in perspective does not compromise bodily self-consciousness. We can, for instance, adopt a third-person perspective in space by imagining what other people can see from their viewpoint (Vogeley et al. 2004). We then refer to egocentric (first-person) or allocentric (third-person) reference frames. Different reference frames can naturally be used in spatial navigation (Maguire et al. 1998), simulation of actions (Ruby & Decety 2001) or episodic memory (Freton et al. 2014). There is a natural tendency for memories to be recalled from the first-person perspective (Freton et al. 2014).

Vogeley and Fink define a set of regions associated with the first-person perspective in a variety of tasks (Figure 4), namely the right inferior parietal cortex, the medial parietal cortex and the medial prefrontal cortex (Vogeley & Fink 2003). This network of regions closely resembles the default network.

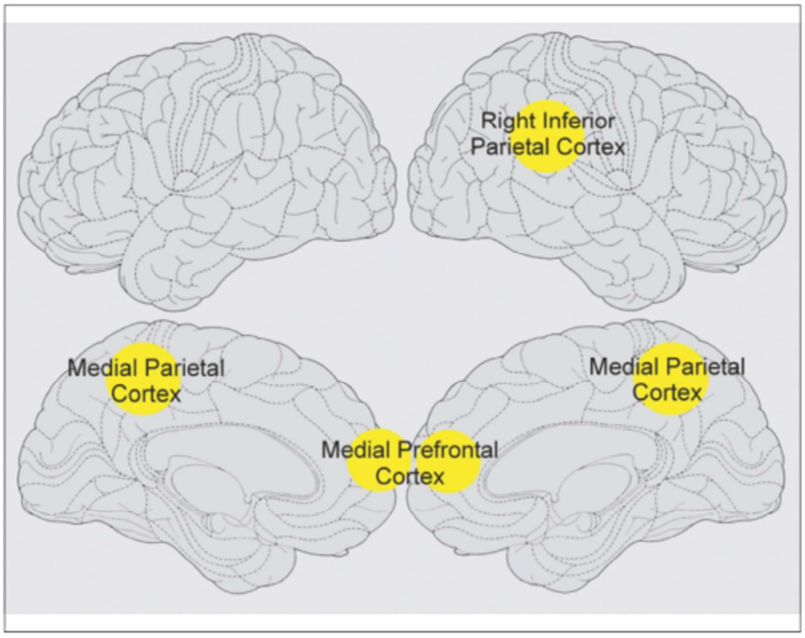


Figure 4

Regions of activation during first-person-perspective tasks. Medial cortical regions, that comprise anterior medial prefrontal, medial parietal and posterior cingulate cortex, are hypothetically recruited if such a state of 'core self' is instantiated. The right inferior parietal cortex is the implementation site of the body representation, which most probably is involved in the computation of the egocentric reference frame. (figure and legend from (Vogeley & Fink 2003)**)**

The "right inferior parietal cortex" (Figure 4) can correspond to regions with different names in the literature: temporoparietal junction, inferior parietal lobule, angular gyrus, posterior superior temporal sulcus or Brodmann areas 39 or 40 (Bzdok et al. 2013, Caspers et al. 2006). These structures are indeed commonly activated in studies involving perspective taking, such as out-of-body experience (Blanke et al. 2002), full-body illusions (Ionta et al. 2011b), mental own-body imagery (Blanke 2005), egocentric or allocentric spatial strategies in navigation (Boccia et al. 2014, Maguire et al. 1998), memory (Ciaramelli et al. 2010), perspective taking for action (Ruby & Decety 2001), or visual perspective taking (Vogeley et al. 2004). However, even though the localizations are close, the results of these studies do not consistently overlap (Figure 5). The fact that this is a large cortical region, with different cytoarchitectonic areas which are highly variable between individuals (Caspers et al. 2006), could potentially explain the discrepancies in the localizations of effects. It remains an open question as to whether there is a region responsible for perspective taking in these different tasks. Specific meta-analyses should be performed, to assess the actual degree of overlap.

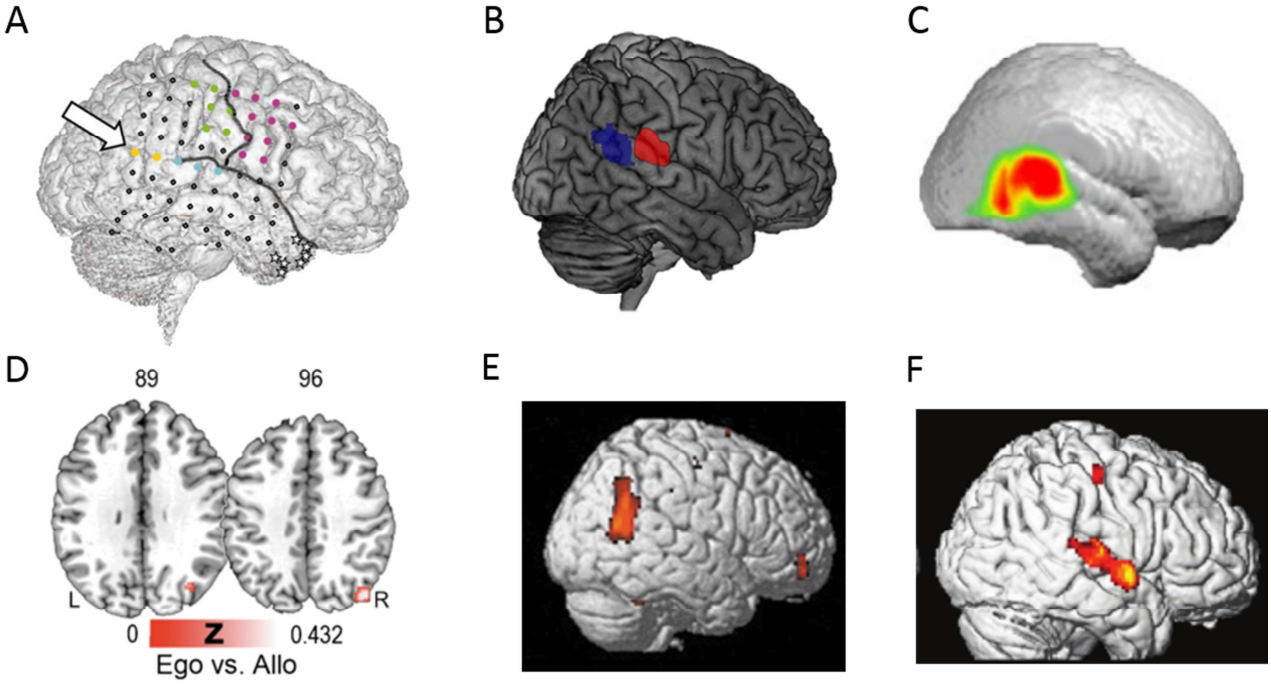


Figure 5

Results of studies involving perspective taking. Only figures showing right inferior parietal cortex activations are included. Reported coordinates are in MNI space, and were transformed from Talairach space when necessary.

(A) Location of subdural electrodes implanted in the brain of an epileptic patient. Focal electrical stimulation on the yellow electrodes (indicated by the arrow) induced out-of-body experiences. No coordinates reported. (Blanke et al. 2002)

(B) Comparison between the area reflecting experimentally induced changes in self-location in healthy participants (red) and the area reflecting pathologically induced changes in self-location in patients with out-of-body experiences (blue). Coordinates of the red cluster: 55 -28 16. (Ionta et al. 2011b)

(C) Activation of the right temporoparietal junction when participants imagine themselves in a certain position and visual perspective. Coordinates of the main cluster: 64 -39 20. (Blanke 2005)

(D) Areas showing higher activation for egocentric (Ego) than allocentric (Allo) spatial strategies. No coordinates reported for the angular gyrus. (Boccia et al. 2014)

(E) Brain areas activated by third-person action simulation compared to first-person. Coordinates of the inferior parietal lobe: 50 -58 30. (Ruby & Decety 2001)

(F) Brain areas activated for first-person visual perspective taking compared to third-person. (Vogeley et al. 2004)

The “medial parietal cortex” of Vogeley and Fink seems to include the dorsal and ventral precuneus as well as posterior/mid-cingulate regions or the paracentral lobule, which are different functional regions (Beckmann et al. 2009, Bzdok et al. 2015, Cavanna & Trimble 2006). The medial parietal cortex is less often associated with out-of-body experiences (but see (De Ridder et al. 2007)) or full-body illusions (but see (Guterstam et al. 2015)), but has been implicated in visual perspective taking (Vogeley et al. 2004), perspective taking for action (Ruby & Decety 2001) and perspective taking in autobiographical memory (Freton et al. 2014). The spontaneous tendency to recall memories from the first-person perspective, for instance, correlates with the volume of the precuneus (Freton et al. 2014). A lesion in the medial parietal regions can lead to egocentric disorientation, which is an impairment in the use of the egocentric reference frame (Wilson et al. 2005). A meta-analysis of the human posterior medial cortex, comparing different postero-medial sub-regions (Bzdok et al. 2015), confirmed previous hypotheses (Burgess et al. 2001, Gramann et al. 2010, Vann et al. 2009) suggesting that switching between egocentric and allocentric reference frames in both memory and spatial domains was specifically performed in the retrosplenial cortex (a region supposedly more ventral than the medial parietal cortex of Vogeley and Fink). Hence, it seems that the retrosplenial cortex indeed implements perspective taking in a variety of tasks. However, the precise mechanisms remain to be characterized, since some studies report more activation for non first-person perspective conditions (Ruby & Decety 2001, Vogeley et al. 2004) while others find more activation for egocentric as compared to allocentric spatial navigation (Boccia et al. 2014).

Conclusion

First-person perspective is therefore very linked to body ownership and self-location. The first-person perspective is very frequently adopted in autobiographical memory, but a third-person perspective can also be adopted without inducing changes in bodily self-consciousness. It can be argued though, that this voluntary shift of perspective still conserves the first-person perspective in the background.

Perspective taking is associated with two regions belonging to the default-network. The medial parietal cortex has a fundamental role in the processing of egocentric and allocentric reference frames. The medial prefrontal cortex seems less

associated with bodily aspects of perspective taking (it is not associated with out-of-body experiences or full-body illusions) and more associated with cognitive ones (theory of mind or autobiographical memory, for example (Spreng et al. 2009)).

3. Agency

Gallagher defines agency as "the sense that I am the one who is causing or generating an action" (Gallagher 2000). Agency unifies bodily self-consciousness and adds a motor/functional component to it (Tsakiris et al. 2007b). Illusions of body ownership have no effect on motor responses (Kammers et al. 2009), but a moving rubber hand can elicit the rubber hand illusion (Kalckert & Ehrsson 2014).

Most often, we feel agency over our actions, such that in ambiguous cases there is a bias toward self-attributing actions (Daprati et al. 1997). But sometimes, there might be a mismatch between the intended and the actual outcome of our actions, so that the outcome is not considered as resulting from our intentional action. In the "comparator model" of agency (Frith et al. 2000), motor commands are compared with the sensory consequences of the action. If there is a mismatch between the two, then non-agency is signaled, in retrospect. The sense of agency has also a prospective component, involving the intention to act and the prediction of action outcomes (Haggard 2017).

The role of the angular gyrus / temporoparietal junction has been consistently related to the signaling of non-agency (Haggard 2017, Sperduti et al. 2011), in both retrospective (Farrer & Frith 2002, Farrer et al. 2003) and prospective (Chambon et al. 2013) accounts of the sense of agency. The insula has in turn shown to be the only structure activated for self-agency (Sperduti et al. 2011).

Therefore, the brain is able to detect non-agency as well as agency, thereby preventing misattributions of actions to the self and distinguishing self- and other-caused actions. In pathological cases, though, the sense of agency can be altered, either by excessively self-attributing actions or by reducing the sense of agency over actions (Haggard 2017).

Conclusion of B: body-image and body-schema

We addressed bodily self-consciousness from different angles: body-ownership, self-location, first-person perspective and agency. Body-ownership and self-location showed that explicit bodily self-consciousness is rather malleable. It can integrate other limbs in the representation of the body and it can be dramatically changed in the case of the full-body illusion. Perspective taking and agency are implemented in a way that can distinguish self and other. Healthy subjects can naturally adopt a third-person perspective and can identify an action that was not self-caused, without compromising the consciousness of their body.

Body-ownership and self-location are part of the *body image*, "a system of experiences, attitudes, and beliefs where the object of such intentional states is one's own body" (Gallagher & Zahavi 2008). Spatial perspective taking and agency relate in turn to the two aspects of *body schema*, "(1) the close-to-automatic system of processes that constantly regulates posture and movement to serve intentional action; and (2) our pre-reflective and non-objectifying body-awareness [...]. The body schema [...] includes our pre-reflective, proprioceptive awareness of our bodily action". At first glance, body-ownership and self-location seemed to be the most fundamental aspects of the self. However, according to this body-schema/body-image theory – originally developed by Merleau-Ponty, it is quite the opposite: "To the extent that one does become explicitly aware of one's own body in terms of monitoring or directing perceptual attention to limb position, movement, posture, pleasure, pain, kinaesthetic experience, and so on, such awareness constitutes aspects of a body image and presupposes the tacit contribution of the body schema" (Gallagher & Zahavi 2008). In other words, the body-image involves the conscious awareness of the body which requires the pre-reflective body-schema.

We will now develop the distinction between the experiential body and the introspective aspect of bodily self-consciousness.

C. A distinction between experience and introspection relative to the body

Let us consider for instance the experiment by Ehrsson and colleagues (Ehrsson et al. 2004), where participants underwent the rubber hand illusion, without any other specific task, before and during fMRI recordings. In this case, participants might simultaneously engage as the subject of experience and the object of introspection, because they experience the illusion and reflect on how it feels. Therefore, the corresponding neuroimaging results (contrasting for instance synchronous vs asynchronous stroking) may correspond to (1) differential experiences, to (2) reflection on the fact experiences differ or to (3) a mixture of both.

Going back to Legrand and Ruby's framework (Legrand & Ruby 2009), the feeling of one's body (option 2) is not self-specific because it does not meet the criterion of non-contingency. For instance, deafferented patients can still make the self/non-self distinction, despite the total loss of proprioception.

Self-specificity is found in actions where sensorimotor loops link efferent signals (motor commands) to their afferent consequences (sensory signals arising from the execution of an action) (Legrand & Ruby 2009). *Re*afferent signals are intrinsically self-specifying because they are the result of efferent signals, while *ex*afferent signals are the result of environmental events (Christoff et al. 2011). More than self-attributing the contents of the action (which is the main point of comparator models of agency), self-specificity of action relies on sensorimotor loops that define the subject as the agent and thereby distinguish between self-specific from non-self-specific information (Christoff et al. 2011, Legrand & Ruby 2009). As Legrand points out, "the crucial difference between perceiving oneself and perceiving others is not purely sensory but sensory-motor: what is self-specific is not a multisensory redundancy but a sensori-motor coherence" (Legrand 2007). These authors see sensorimotor activations as revealing the self-specifying sensorimotor loops. It was indeed shown that corticospinal excitability is associated with the feeling of agency (Weiss et al. 2014). This was considered a low-level sensorimotor marker of agency, on which agency reports could be based.

The discrepancy between the interpretative level and the experiential level was developed by Synofzik, through the notions of Feeling of Agency (FoA) and Judgment of Agency (JoA) (Synofzik et al. 2008). The FoA refers to the non-conceptual, implicit

and low-level feeling of being the agent of an action. The JoA is conceptual, explicit, and refers to the judgement of being the agent of the action. Usually, we do not need a JoA, because ambiguous situations are rare in real life. To experimentally manipulate and dissociate the FoA and the JoA we need implicit measures of agency, reflecting uniquely the experience or the feeling of agency without requiring participants to judge their agency. Many paradigms of agency indeed require subjects to *evaluate* whether they are responsible for the movement or not (Farrer et al. 2003), which introduces the JoA on top of the FoA. Haggard and colleagues created the intentional binding paradigm (Haggard et al. 2002), where participants perform an active or a passive finger movement, followed by the presentation of a tone. When the movement is voluntary, there is an attraction of the two events (movement and tone) in time (intentional binding). The tone appears as a consequence of the movement; triggered by the participant's action. The evaluation of the time distance between the events is therefore considered a proxy for the feeling of agency. The SMA proper has been showed to reflect intentional binding (Kühn et al. 2013), as well as the pre-SMA in non-invasive brain stimulation experiments (Cavazzana et al. 2015, Moore et al. 2010). Other neuroimaging experiments directly compared FoA and JoA, and showed that the pre-SMA was associated with the feeling of being in control (as well as the rostral cingulate zone and the dorsal striatum), whereas the feeling of being out of control corresponded to TPJ activations. Explicitly judging agency was associated with anterior prefrontal cortex activations (Miele et al. 2011). These results support a dissociation between the JoA and the FoA, each being associated with specific brain structures.

De Vignemont also applies the distinction between feeling and judgment to ownership of the body (de Vignemont 2011). Apart from the natural ability to make judgments about our own body, "there is something it is like to experience parts of my body as my own, some kind of non-conceptual intuitive awareness of ownership" (de Vignemont 2011). Experimental literature on body ownership does not usually make this distinction. In the rubber hand illusion paradigm, hand ownership is often assessed via questionnaires that may tap more into the judgment of ownership.

Conclusion

We have seen here how we can distinguish between the embodied pre-reflective first-person perspective and the explicit experience of the body. This dissociation is reminiscent of the dissociation we discussed earlier, between the "I" and the "Me". In our view, both feeling of agency and feeling of ownership relate to the "I", to the experiencing subject. On the contrary, judgment of agency and judgment of ownership relate to the "Me", to the introspection of the (bodily) self.

We will now see how some of these bodily processes relate to higher-order forms of self.

D. Interactions between bodily processes and higher order self

Interestingly, among the brain regions associated with the bodily aspects we mentioned in part III.B, the default-network seems to be often relevant. If we think in terms of overlapping regions, we could imagine that these bodily processes are somehow linked to the self-related processing we discussed in II.B, that often involves these same regions. However, direct evidence remains scarce.

The case of the enfacement illusion demonstrates that the representation of one's own face can be modulated by visuo-tactile stimulations. This is an example of how bodily factors can influence higher order aspects of the self, because one's own face is an important representation of the (bodily) self, probably more based on memory than on online bodily monitoring (I cannot see myself but I know what I look like from the times I looked at myself in a mirror). The own face is part of one's identity. For Tsakiris (Tsakiris 2010), even the case of the rubber hand illusion reveals that multisensory processes interact with higher order representations of the body. The visual shape of the object has to match a pre-existing hand model in order to be embodied. Thus multisensory mechanisms also have to interact with stored body models.

Autobiographical memory and self-projection in the future are very much linked to perspective, since these processes can occur from a first- (when remembering or imagining oneself from inside the body) or third-person perspective

(when adopting another person's perspective). Autobiographical memories are better recalled when a posture congruent with the recalled event is adopted (Dijkstra et al. 2007). The body-memory link can not only happen at the level of retrieval of memories but also at the moment of encoding. Bergouignan showed that recollection of episodic memories is impaired when episodes are encoded under the out-of-body illusion, rather than in-body, suggesting that efficient episodic memory encoding requires an embodied first-person perspective on the world (Bergouignan et al. 2014).

Conclusion

We saw here how bodily signals can shape and interact with higher order forms of self, such as self-representation and autobiographical memory. In the experiments we mentioned here, the body is manipulated in an explicit way: either through face stimulation, body postures or full-body illusions. However, we can think of a much more implicit and continuous role of the body, which could underlie lower-level, pre-reflective aspects of the self (not impaired in the experiments reviewed). This led us to look for other kinds of signals, which would be pervasive and not necessarily noticeable: visceral signals.

E. From the somatosensory/motor body to the visceral body

Body ownership, self-location, first-person perspective and agency are important features of bodily self-consciousness and involve mostly proprioceptive and somatosensory/motor signals. These signals seem to contribute to higher order levels of self, but other kinds of bodily signals may play an important role as well.

Indeed, when participants experience full-body illusions or when patients have out-of-body experiences, a sense of explicit bodily self is disturbed while something more basic about self-consciousness is still preserved. Participants are still able to say "I feel something", they still have this pre-reflective form of self. The multisensory integration leading to these illusions is integrating misleading signals, however there may be other bodily signals responsible for the preserved "I".

These could be signals from visceral organs, namely the heart and the gut, which are of extreme importance to the organism. Visceral signals are relayed through redundant pathways and target multiple cortical brain regions, ensuring the transmission and broad availability of these signals. The constant communication between the brain and the internal organs is necessary for the regulation and maintenance of physiological parameters, and thus for homeostasis (Damasio 1999).

Contrarily to somatosensory/motor signals, visceral signals are still relayed to the brain in locked-in patients (Park & Tallon-Baudry 2014). These patients are paralyzed but are still conscious and maintain a sense of self (in a recent survey, 72% of the patients reported being happy (Bruno et al. 2011)). Damasio describes the case of a patient with asomatognosia who does not sense her somatosensory body but whose sense of self is preserved: "a patient who had a temporary loss of the sense of her entire body frame and body boundary (both left and right sides) but was nonetheless well aware of her visceral functions (breathing, heartbeat, digestion) and who could characterize her condition as a disquieting loss of part of her body but not of her "being" " (Damasio 1994). From this case, Damasio hypothesizes that "some body representations may be of greater value than others to ground the mind, namely, those that pertain to the organism's interior, specifically to the viscera and internal milieu" (Damasio 2003). Accordingly, Damasio interprets the epigastric auras preceding an epileptic seizure as a disruption of the brain mapping of internal bodily states, causing the loss of consciousness (Damasio 2003).

How could visceral signals be involved in the implementation of the self?

IV. The visceral body

A. Theoretical considerations about the visceral body and the self

1. A bodily-centered reference frame for the self

We discussed in part III the importance of the body for the self and, in particular, of somatosensory/motor mechanisms. However, even if these bodily signals are constantly integrated by the brain, the signals we mentioned were mostly responses to a stimulus or feedback from certain actions. In our view, the *continuous* monitoring of the body might constitute the fundamental ground upon which selfhood is based.

The monitoring of the body includes somatosensory and motor signals, but also includes visceral signals, which as we saw might be of a greater importance. **The main proposal of this thesis is that the integration of these signals in the brain may constitute a bodily-centered reference frame that would define the self. Any self-related process would be anchored to this reference frame, in order to be labeled as subjective, by the brain. We evoked before that rather than looking for brain regions implementing the self, one could look for a *mechanism*. We propose that the brain monitoring of bodily signals might be this mechanism, taking place in any brain region where a self/non-self distinction is relevant.**

Until now, we had a loose definition of the self: it included higher-order cognitive forms of self (autobiography, self-recognition, introspection…), but also a more basic form of self, the first-person perspective of the experiencing or acting self. We are now talking of a very low-level form of self: the organism as a unified entity. This low-level (unconscious, or pre-reflective) self is very much related to what we previously defined as the "I" and could underlie any other (higher-order) form of self.

This idea of a *neural subjective frame* was developed in particular by Park and Tallon-Baudry (Park & Tallon-Baudry 2014), and was experimentally tested in the context of visual perception (we will return to this later). The experimental purpose of this thesis was to directly address this theory by tackling low- and higher-level forms of self.

2. *About visceral signals*

The neural subjective frame theory is very much inspired by Damasio's theories of the self (Damasio 1994, 2003). Both highlight the importance of bottom-up signals, from the viscera to the brain. However, for Damasio, the variations of these signals carry the important information: "I believe subjectivity depends in great part on the changes that take place in the body state during and after the processing of object X" (Damasio 1994). On the contrary, for Park and Tallon-Baudry, the content of the signals (their intensity, for instance) is not necessarily required for the bodily-centered reference frame. The simple monitoring of these signals would be sufficient. Regardless of what they say about the state of the body, these signals say that a body is there, and that would be enough to anchor subjective processes. If these visceral signals indicate that important changes are taking place, then of course this information might be taken into account.

Importantly, this hypothesis can be tested. If we find a link between the self and viscera-brain coupling, is it associated with variations of physiological signals? We tried to address this question in our experimental work.

We will now focus on the visceral body. We will describe the pathways from the viscera to the brain and see how viscera-brain coupling can be studied and how it can contribute to cognitive processes.

B. Pathways from the viscera to the brain

Signals about the state of the organs in the thorax, abdomen and pelvis are conveyed to the brain via visceral afferent neurons. They convert mechanical or chemical changes into electrical information that they transmit to the brain via vagal (parasympathetic) or spinal (sympathetic) pathways (Critchley & Harrison 2013, Jänig 1996, Saper 2002, Vaitl 1996) (Figure 6).

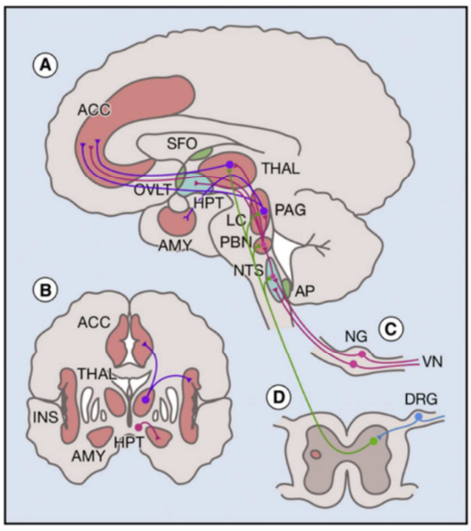


Figure 6

Diagram of Viscerosensory Paths and Centers in the Human Brain. Depicted schematically are (A) parasagittal, (B) coronal brain sections, (C) nodose ganglion of vagus nerve, and (D) section of the spinal cord. These figures illustrate viscerosensory centers and interoceptive neural pathways. Visceral afferent inputs with cell bodies in the dorsal root ganglion (DRG) enter the spinal cord (lamina I) and ascend in the spinothalamic tract (light green) to terminate in viscerosensory thalamus (THAL) with earlier outputs to the nucleus of the solitary tract (NTS), parabrachial nucleus (PB), and periaqueductal gray matter (PAG). Viscerosensory inputs carried by the vagus nerve (VN) with cell bodies in vagus nerve ganglia (nodose ganglion, NG) terminate in the NTS and then pass to PB, PAG and THAL (pink). Information is relayed from THAL, PAG and PB to hypothalamus (HPT), amygdala (AMY), anterior cingulate cortex (ACC), and insula (INS), the latter being the primary site of viscerosensory cortical representation. (figure and legend from (Critchley & Harrison 2013)**)**

85% of the fibers of the vagus nerve are visceral afferent fibers. They project viscerotopically to the nucleus of the solitary tract (NTS), where visceral information converges (Jänig 1996). Visceral signals are then conveyed to the parabrachial nucleus which dispatches the signals to hypothalamus, thalamus and the cortex. This pathway carries mostly *motivational information*, i.e. hunger, satiety, thirst, nausea, and respiratory sensations (Critchley & Harrison 2013).

In contrast, only 1.5-2% of all spinal afferents are visceral, the vast majority being somatic (from the skin, joints, tendons and muscles). Spinal afferent neurons project primarily to lamina I and V of the dorsal horn of the spinal cord. Spinal neurons that are excited by spinal visceral afferents are also excited by somatic afferents, thus leading to viscero-somatic convergence (Jänig 1996). These signals, which mainly transmit information about tissue damage (Critchley & Harrison 2013), are then carried centrally by the spinothalamic tract. This pathway also projects to

the NTS and parabrachial nucleus, where these signals converge with the vagal pathway (Saper 2002).

The amygdala, the ventral anterior cingulate cortex and the posterior insula are considered the direct targets of visceral information (Armour & Ardell 2004, Craig 2002, Critchley & Harrison 2013, Saper 2002). Visceral information is also mapped in somatosensory cortices SI and SII (Damasio & Carvalho 2013).

A recent meta-analysis combined the results from a diversity of studies focusing on the central autonomic processing (Beissner et al. 2013) (based on the high-frequency component of heart rate variability and electrodermal activity). This meta-analysis evidenced a central autonomic network, whose core involves the left amygdala, the right anterior and left posterior insula, and midcingulate cortices (Figure 7). Whether these structures are receiving visceral information (bottom-up) and/or sending signals back to the viscera (top-down) is not specified in the meta-analysis.

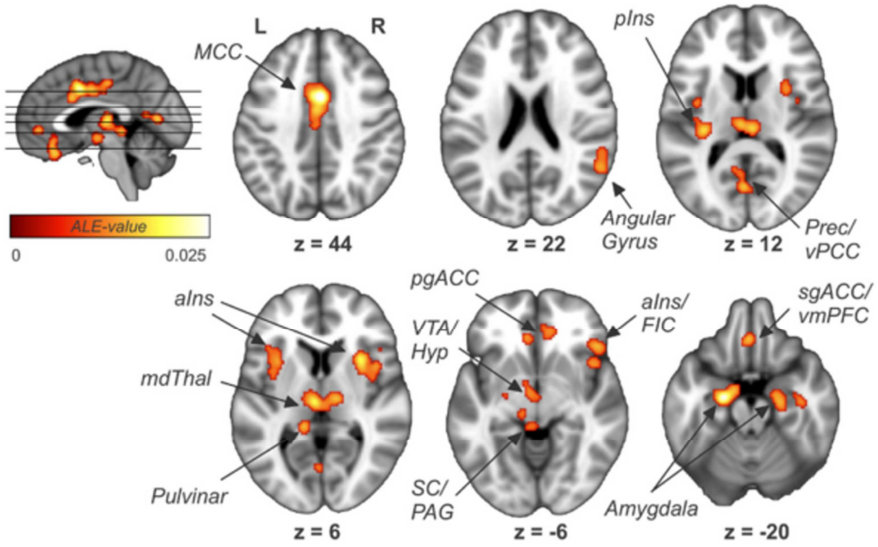


Figure 7

Results of the pooled analyses of all studies showing general brain regions involved in autonomic processing. Prec, Precuneus; vPCC, ventral posterior cingulate cortex; mdThal, mediodorsal thalamus; pgAcc, pregenual ACC; VTA, ventral tegmental area; Hyp, hypothalamus; SC, superior colliculus; PAG, periaqueductal gray; FIC, frontoinsular cortex; L, left; R, right. (figure and legend from (Beissner et al. 2013)**)**

The viscera to brain pathways are the bottom-up branch of the organ regulation loops, which ensure homeostasis. Interestingly, as Vaitl notes (Vaitl 1996), the number of afferent fibers is significantly larger than the number of efferent fibers, which suggests that a bottom-up transfer of information is prioritized relative to top-down transfer of information. While mechanical changes can be quickly detected, chemical changes occur on a slower time scale, which implies that the brain is being informed about the state of the body at different time scales (Park & Tallon-Baudry 2014). The fact that this bottom-up information reaches a wide variety of high-level cortical areas suggests that visceral information is widely available in the brain. It could thus interfere with ongoing cortical activity and potentially with a number of cognitive processes, thereby playing a role beyond homeostatic regulation.

C. Resting state cortical activity and physiological signals

Physiological signals have long been considered a source of noise affecting neuroimaging techniques (Birn 2012, Glover et al. 2000, Shmueli et al. 2007) and were therefore systematically removed during data analysis. However, there is now evidence that these signals are not mere noise (Beissner et al. 2013, Iacovella & Hasson 2011, Thayer et al. 2012), notably during the resting state.

Indeed, during the resting state, skin conductance levels (Fan et al. 2012, James et al. 2013, Nagai et al. 2004), respiration fluctuations (Yuan et al. 2013), heart rate (de Munck et al. 2008) and hypoglycemia (Teves et al. 2004) co-vary with BOLD activity in a variety of cortical regions, beyond the insula or the anterior cingulate cortex (mid-cingulate, precuneus, orbitofrontal cortex, for instance). Resting state functional connectivity co-varies with heart rate variability (Chang et al. 2013). Interestingly, regressing out respiratory fluctuations from BOLD signals decreases the correlation between BOLD and alpha EEG power (Yuan et al. 2013), suggesting that there is a neuronal coupling between BOLD, alpha and respiration. This finding is complemented by recent evidence showing a co-variation of alpha, beta and gamma power with fluctuations of arterial CO_2 (Driver et al. 2016).

Whether these results reflect a bottom-up influence of physiological signals on brain activity or a top-down regulatory influence of the brain on physiological signals

is not known. During my PhD, I contributed to a study showing that the gut had a bottom-up influence on brain resting state activity: the phase of the gastric basal rhythm modulates the amplitude of the alpha rhythm, in the right anterior insula and occipito-parietal regions (Richter et al. 2017).

Taken together, these results show that physiological information can reach many different cortical structures and that resting state brain dynamics are coupled with physiological fluctuations.

D. Heart and brain

From the perspective of the body-centered reference frame hypothesis, we are interested in visceral signals that are emitted by an internal organ and relayed to the brain. Cardiac signals are good candidates, because they are emitted by the heart at each contraction (60-100 bpm) and because they are relayed up to the brain where they elicit measurable heartbeat-evoked responses.

1. Stimuli processing and the timing of the cardiac cycle

Several results show that the perception of a stimulus can be modulated by the phase of the cardiac cycle (systole or diastole) at which the stimulus is presented (for a review (Park & Tallon-Baudry 2014)). Microsaccades are more frequent in the early phase after the R-peak supporting the existence of a coupling between heartbeats and the oculomotor system (Ohl et al. 2016). The reaction time to auditory, visual and tactile stimuli detection increases when stimuli are presented around the R-peak (Edwards et al. 2007, Saari & Pappas 1976). Also, when arterial pressure is low, the auditory N1 and the visual P1 components of sensory potentials are larger in amplitude. More recently, it was shown that stimuli presented synchronously with heartbeats are less likely to be consciously perceived (Salomon et al. 2016).

In addition, cardiac phase is also thought to interfere with the processing of threat-related stimuli (Garfinkel & Critchley 2015). Indeed, fearful faces presented at systole are perceived as being more intense and enhance amygdala activity relative to faces presented during diastole (Garfinkel et al. 2014). This effect is not observed for

happy or neutral faces, and is also absent or at least weaker for disgusted faces. Race-threat stereotypes elicited during systole are more likely to activate racial biases than when presented during diastole (Azevedo et al. 2017). In contrast, painful stimulation during systole leads to decreased pain perception (Edwards et al. 2002), decreased amygdala activity but increased insular activity (Gray et al. 2009). Depending on the phase of the cardiac cycle, different stimuli are prioritized. During systole, the processing of potentially threatening stimuli (fearful faces, racial biases) is enhanced relative to the processing of other kinds of stimuli (visual, painful…), in order to allow the quick and appropriate selection of behavioral responses, between fight and flight (Garfinkel & Critchley 2015).

2. Heartbeat-evoked responses

The neural monitoring of the heart can be studied by measuring heartbeat-evoked responses (HERs) (Schandry et al. 1986). HERs correspond to the brain activity, recorded with electroencephalography (EEG) or magnetoencephalography (MEG), *evoked* by each heartbeat, i.e. *locked* to each heartbeat (Figure 8).

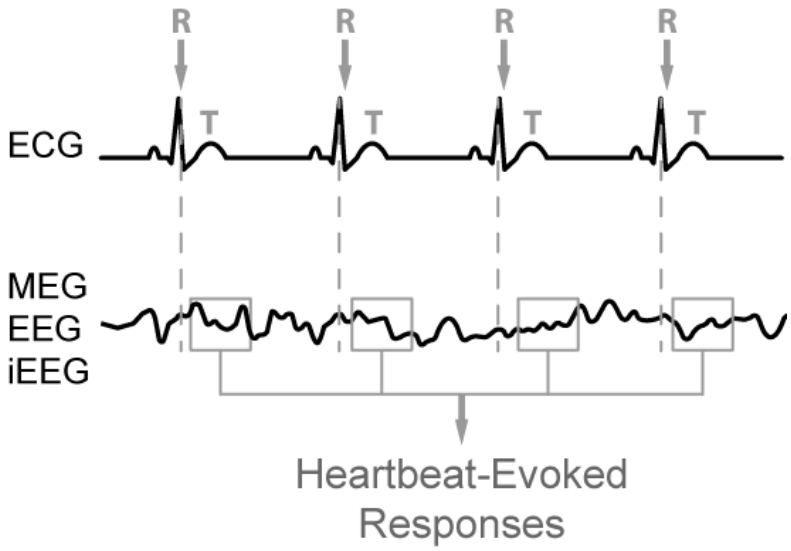


Figure 8

What is a heartbeat-evoked response (HER)? Brain activity acquired with MEG (magnetoencephalography), EEG (electroencephalography) or iEEG (intracranial EEG) is extracted and averaged, locked to a certain phase of the cardiac cycle (here the R-peak of the ECG (electrocardiogram), but it could be the T-peak as well). The R-peak is the largest peak of the ECG and corresponds to the depolarization of the ventricles, which creates the contraction of the heart (systole). The T-wave corresponds to ventricular relaxation (diastole).

a) Origin of the cardiac signal that reaches the brain

Information about heart contraction is primarily acquired via mechanoreceptors, i.e. neurons sensing changes in pressure following blood ejection. The heart walls are densely innervated by these mechanoreceptors, which can have different conduction velocities (depending on whether they are myelinated or not) (Shepherd 1985). The inner curvature of the aorta also contains mechanoreceptors (baroreceptors), which transduce very precise information about changes in the aortic diameter (Armour & Ardell 2004, Garfinkel & Critchley 2015). During diastole, these mechanoreceptors are silent, but when blood is ejected, during systole, they generate bursts of activity that are relayed through vagal and spinal pathways as described above to the dorsomedial part of the NTS. In the rat, it has been shown that NTS neurons are indeed responsive to stimulation of baroreceptor afferents (Nosaka et al. 1995). After the NTS, this signal could be transmitted to the parabrachial nucleus and to sub-cortical and cortical areas, where it can be detected in the form of HERs, with EEG or MEG.

It should be noted that somatosensory pathways are likely to convey heartbeat-related information as well. Indeed, the chest is innervated by somatosensory afferents that might respond to the impact of the heart on the chest, at each heartbeat (Khalsa et al. 2009).

The precise origin of HERs is not well known. Invasive experiments would be needed to answer this question. At present, the baroreceptor hypothesis is the most widely reported.

b) Characterization of the HER waveform

In EEG studies, HERs were first described as a broad positive wave, over frontal electrodes (Schandry & Montoya 1996), however later studies showed the existence of HERs in parietal and central regions (Kern et al., 2013, for a review). The first study exploring HER sources (Pollatos et al. 2005a) used the dipole source localization technique and found dipole locations compatible with the anterior cingulate cortex, the medial frontal gyrus, the right insula and the left somatosensory cortex. More precise techniques of source localization (minimum norm estimation, MNE) based on MEG recordings, which have a better spatial resolution than EEG, have also shown the involvement of the right inferior parietal lobule and ventral

anterior cingulate cortex in the generation of HERs (Park et al. 2014). HERs were also found in the insula, but with a region of interest analysis of high-density EEG data (Couto et al. 2015).

Kern and colleagues (Kern et al. 2013) characterized HERs obtained with intracranial recordings from electrocorticography (ECoG), with a grid placed over the left primary somatosensory cortex, during resting state. They showed that, in this region, HERs are a biphasic potential, with a positive peak at 280ms and a negative peak at 360ms after the R-peak (Figure 9).

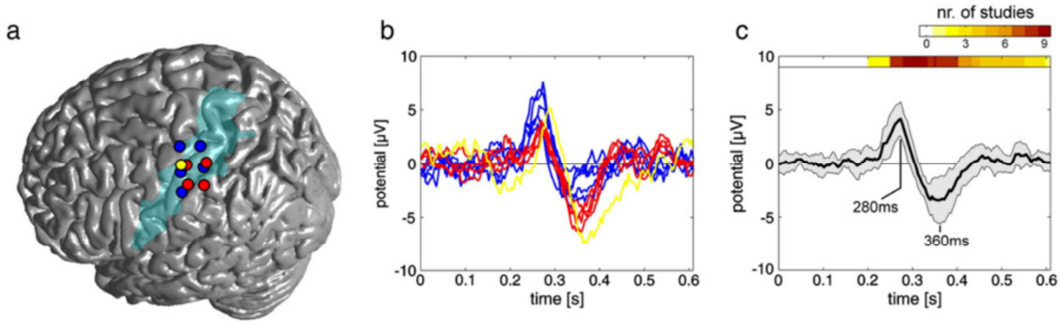


Figure 9

The delayed biphasic heart cycle-related potential. On the left side (a), electrode contacts that recorded the biphasic potential (represented as colored circles) are shown superimposed on a standard brain surface; blue, red and yellow circles correspond to electrode contacts of P1, P2 and P3, respectively (in P3, the electrode position was mirrored to the left hemisphere for better inter-individual comparison); the transparent blue area indicates the left somatosensory cortex. b) Time course of all biphasic potentials at the electrode contacts shown in (a); color red as in (a); c) Grand average over all selected electrodes; the black curve represents the average potential as the mean over electrodes; the corresponding standard deviation is coded as gray band. The HER latency of 12 EEG previous studies is shown at the top of the graph. The absolute number of studies reporting HER occurrence at the corresponding point in time is color coded. The latency of the biphasic heart cycle-related potential is in good accordance with the existing literature about HER latency in EEG recordings. (figure and legend adapted from (Kern et al. 2013)**).**

Many studies reported different effect latencies, ranging from 200 to 650ms after the R-peak (Kern et al. 2013). Based on (Fagius & Wallin 1980), Gray reports that, in humans, baroreceptor information should reach central visceral centers around 400-800ms after the R-wave of the electrocardiogram (Gray et al. 2007), which is inconsistent with some of the results of the literature on HERs. Schandry and Weitkunat (Schandry & Weitkunat 1990) estimate, in turn, that HER latency

should be 280-370ms after the R-peak, considering that the pressure peak in the heart and aorta occurs at 180 and 220ms respectively.

Therefore, HERs do not seem to be a standardized waveform, with a characteristic location or latency.

c) HER amplitude modulation by different cognitive factors

Since the first paper by Schandry and colleagues (Schandry et al. 1986) and in particular since the year 2010, a number of studies have explored the relationship between HER amplitude and different kinds of cognitive processing. Most of these studies have focused on cardiac interoception, i.e. the ability to perceive one's own heartbeats. In brief, HER amplitude is larger when one pays attention to heartbeats and also for good heartbeat perceivers. The link between HERs and interoception will be presented in more detail in the next section on cardiac interoception. Here, we will focus on other cognitive effects on HER amplitude. Indeed, the brain responds to heartbeats even when one is not paying attention to them.

(1) Emotional processing

The negativity observed at 250-430ms after the R-peak over frontocentral EEG sensors was enhanced during an affective judgment task compared to a physical judgment task (Fukushima et al. 2011). HER amplitude correlated with self-rated empathy scores, which led the authors to propose that the neural monitoring function could be involved in empathy processing. However, judging whether a person feels "positive" or "negative" based on a picture of the eyes might relate more to a valence judgment task than to an empathy task. No comparison between positive and negative conditions was performed in this study, but another study targeted this aspect (Couto et al. 2015). Participants were presented with videos inducing positive, negative and neutral emotional states. A region of interest analysis of high-density EEG data showed that HER amplitude was modulated over a fronto-insulo-temporal network, comprising the insula, the anterior cingulate cortex and the lateral orbitofrontal cortex.

(2) Arousal and sleep

The induced arousal state can have an impact on HER amplitude. The size of the HER effect correlates with ongoing alpha power (Luft & Bhattacharya 2015). More general arousal levels were studied by comparing HERs during wake and the

different sleep stages (Lechinger et al. 2015). The frontal HER positivity, around 300-450ms, decreases from wake to deep sleep and increases again during rapid eye movement sleep.

(3) Visual perception

The amplitude of HERs in the ventromedial prefrontal cortex and the right posterior inferior parietal lobule can predict whether a faint visual stimulus will be consciously perceived or not (Park et al. 2014). This effect was significant at around 150ms after the T-peak and corresponds to changes in visual sensitivity (d'), not to changes in the criterion of response. Importantly, no difference was observed in the cardiorespiratory measures acquired (heart rate, heart rate variability, peripheral blood pressure, respiratory patterns).

(4) Pain, stress and other disorders

Stress can modulate HER amplitude (Gray et al. 2007), and pain leads to HER suppression (Shao et al. 2011). A link was found between HER and diabetic neuropathy (Leopold & Schandry 2001), 3D cognitive fatigue (Park et al. 2015), food deprivation (Schulz et al. 2015a) as well as a number of personality factors, such as depression (Terhaar et al. 2012) or depersonalization/derealization disorder (Schulz et al. 2015b).

(5) Bodily self

In the last year (2016), two papers have been published concerning the relationship between the bodily self and HER amplitude. The first study (Park et al. 2016) measured HERs during full-body illusions and showed that HER amplitude differs depending on whether visual and tactile stimulations are synchronous or asynchronous. This effect was associated with the posterior cingulate cortex / supplementary motor area, and was significant around 300ms after the R-peak. Moreover, the amplitude of HERs correlated with the subjective effects of the illusion.

The second paper explored the relationship between HER amplitude and a modified version of the enfacement illusion (Sel et al. 2016). Here there was no tactile stimulation, but a change in the luminosity of the pictures, synchronous or asynchronous relative to the participants' heartbeats. HER amplitude differed between synchronous and asynchronous conditions (centro-parietal electrodes at around 250ms after the R-peak). However, because in the synchronous condition the

HER overlapped with the response to the visual stimulation, it is difficult to determine whether these effects are truly due to HER differences, to differences in the visual response or to an interaction of both.

Conclusion

The variety of results is quite striking. These studies report different topographies and latencies of effects. Also, as we will see later, HER analysis requires careful controls, which are not always applied in these studies. Moreover, these studies target very different aspects of cognition, which all appear related to HER amplitude. From this evidence, it is hard to understand the specificity of HER effects.

We will now focus on cardiac interoception, which is more directly related to the processing of cardiac signals.

3. Cardiac interoception

a) Cardiac interoception measures

Interoception refers to the sensitivity to one's own internal body, as opposed to exteroception or proprioception, which refer to the sensitivity to the external environment or to the position of the body in space, respectively. Even though interoception refers to the *explicit* attentional orientation to internal bodily states, it is also used sometimes to designate any kind of *implicit* visceral process ("interoceptive pathways", for instance, refer to all pathways conveying visceral signals, regardless of whether the signals are consciously perceived or not). Here, we will review data on *explicit* cardiac interoception, i.e. when attention is directed to the heart.

Two major experimental paradigms were developed in order to assess individual abilities to detect heartbeats. In the heartbeat counting task (Schandry 1981), participants have to concentrate on and count their heartbeats, without directly taking their pulse, during short blocks of different duration. In the heartbeat detection task (Kleckner et al. 2015, Whitehead et al. 1977), participants have to report whether tones or lights are presented synchronously or asynchronously relative to their heartbeats.

Interoceptive abilities are highly variable in the general population (Herbert & Pollatos 2012), even in children (Koch & Pollatos 2014). They appear to be a rather stable trait (Critchley & Harrison 2013), and are not improved even with extensive meditation training (Khalsa et al. 2008). Moreover, cardiac awareness correlates with gastrointestinal awareness (Herbert et al. 2012, Whitehead & Drescher 1980), suggesting that there is a general ability to perceive one's own internal states across visceral modalities.

Garfinkel distinguishes interoceptive *abilities* from interoceptive *sensibility* and interoceptive *awareness* (Garfinkel et al. 2015). Interoceptive abilities correspond to the accuracy in heartbeat counting tasks. Interoceptive awareness refers to the subjective evaluation of one's own interoceptive abilities and can be assessed by rating confidence on interoceptive judgments. Interoceptive sensibility refers, in turn, to the way one experiences internal sensations; it can be assessed through self-report questionnaires. These three interoceptive dimensions were proven to be dissociated, with interoceptive accuracy being partly predicted by interoceptive sensitivity and interoceptive awareness.

b) Neural correlates

Using a heartbeat detection paradigm, Critchley and colleagues (Critchley et al. 2004) investigated the neural correlates of cardiac interoception, in terms of attention focus and performance. Orienting attention to heartbeats leads to an increase in the activity of the bilateral insula, right opercular, somatomotor, anterior cingulate/supplementary motor cortices, compared to orienting attention to external auditory tones. Moreover, interoceptive abilities correlated with activity in the right anterior insula / operculum and with the gray matter volume of this region. Other studies confirmed the role of this interoceptive network in attentional orientation to heartbeats (Pollatos et al. 2007b), and in particular the role of the insula (Ronchi et al. 2015, Schulz, Wiebking et al. 2014, Zaki et al. 2012). This suggests that the insula but also the anterior cingulate cortex integrate cardiac signals which can potentially be consciously accessed (Garfinkel et al. 2013b).

However, other studies have questioned the central role of the insula and the anterior cingulate cortex in interoceptive processing, suggesting that somatosensory pathways may be crucial for cardiac interoception (Khalsa et al. 2009).

HERs have also been studied in the context of cardiac interoception tasks. Orienting attention to heartbeats was shown to modulate HER amplitude (Montoya et al. 1993), which also differs between good and poor heartbeat perceivers (Pollatos & Schandry 2004, Pollatos et al. 2005b, Schandry & Montoya 1996, Schandry et al. 1986). The sources corresponding to these effects are the right insula, the anterior cingulate and left secondary somatosensory cortices (Pollatos et al. 2005a, 2016). Activations observed in fMRI in the insula and anterior cingulate cortex could thus be due to differential HERs. Moreover, interoceptive deficits in patients were associated with a decreased HER amplitude modulation (García-Cordero et al. 2016).

In addition, the mechanisms underlying interoceptive awareness in the insula seem to involve enhanced GABA concentration in this structure, which could be responsible for biasing the balance of interoceptive/exteroceptive processing towards interoceptive processing, by inhibiting exteroceptive processing (Wiebking et al. 2014).

c) Relationship between cardiac interoception and cognition

The relationship between interoception and emotion has been a matter of research (Damasio & Carvalho 2013), following the James-Lange theory (James 1884), which states that emotions *are* the feeling of our automatic bodily reactions to emotive stimuli. In accordance with the theory, good heartbeat-perceivers evaluate emotional pictures as being more arousing and have larger P300 components of visual evoked potentials to emotional pictures compared to poor heartbeat perceivers (Herbert et al. 2007; Pollatos et al. 2005b, 2007a). Both interoception and emotional processing involve the anterior insula (Zaki et al. 2012), suggesting a possible link between the two functions. Interoceptive abilities appear also to be altered in alexithymia (Herbert & Pollatos 2012, Shah et al. 2016).

Additionally, high interoceptive accuracy was shown to correlate with better decision making (Dunn et al. 2010, Kandasamy et al. 2016, Werner et al. 2009), enhanced memory (Garfinkel et al. 2013a), attention to visual stimuli (Matthias et al. 2009) and better prospective memory (Umeda et al. 2016). Eating disorders were in turn associated with lower interoceptive abilities (Herbert & Pollatos 2012, Pollatos et al. 2008). Interoceptive abilities are also altered in autism (Dubois et al. 2016) and anxiety (Critchley et al. 2004, Schandry 1981).

d) What cardiac interoception does and does not tell us

Taken together, these results suggest that cortical targets of visceral signals are also involved in the awareness of those signals. However, there are many factors that could explain the link between interoception and cognitive domains. For instance, since heartbeat detection performance correlates with anxiety (Critchley et al. 2004, Schandry 1981) and gender (Critchley et al. 2004), it may be that anxiety or gender, rather than heartbeat detection performance, can explain some of the results relating interoception and emotion. Since a number of factors correlates with heartbeat perception accuracy (age, fitness, gender or body fat (Critchley et al. 2004)), these should be controlled for, especially when testing different populations. Inter-individual correlations should be performed in large samples of participants and the number of trials is also crucial in this type of task (Kleckner et al. 2015). Thus, the causal link between explicit interoception and cognitive factors remains to be demonstrated, which is an undoubtedly challenging achievement in human cognitive neuroscience.

The actual measure of interoceptive awareness is questionable, since the two tasks do not always yield congruent results (Brener & Ring 2016, Phillips et al. 1999, Schulz et al. 2013a). While one requires only an internal focus (heartbeat counting task), the other requires multimodal monitoring of signals with different intensities (heartbeat detection task) (Garfinkel et al. 2015). Also, some results suggest that cardiac interoception does not correlate with respiratory sensitivity (Harver et al. 1993), which either questions the validity of the methods used or challenges the idea that there is one general interoceptive function.

Cardiac interoception tasks were developed to understand the neural correlates of the awareness of visceral signals. However, the origin of these cardiac-related signals is not clear. They could come from the activity of baroreceptors in the aorta wall or from somatosensory afferents from the skin (Brener & Ring 2016). Heartbeat perception could be mediated by the feeling of the heart on the chest or by the feeling of pulsating vessels (which can happen in constrained fMRI settings or when using ECG measuring devices such as finger pulse oximeters, which may exert some pressure on the finger). Khalsa and colleagues showed that the insular pathway is not necessary to perform this task, since signals from the skin to somatosensory cortices are sufficient (Khalsa et al. 2009). The particular case of a patient with an

external heart in parallel to his own also suggests that somatosensory signals coming from the external device guided the patient when performing the heartbeat detection task (Couto et al. 2014). Moreover, people with less body fat (Kleckner et al. 2015) and people with bigger hearts are better at performing the task (Vaitl 1996), possibly because they can better feel the heart beating inside the chest. It is therefore legitimate to ask whether this task can reliably tell us something about the awareness of purely ascending *visceral* signals, since it may rely on a combination of both visceral and somatosensory ascending signals.

Further, the reason why some people perform this task better than others remains unclear. This might be due to differences in cardiac function, such as the size of the heart (Vaitl 1996) or blood pressure levels (O'Brien et al. 1998). More accurate performers may simply have stronger (and thus easier to detect) ascending cardiac signals. Another hypothesis is that it could be linked to individual attentional capacities. Better heartbeat-perceivers could be those who are able to enhance interoceptive awareness over other sensory modalities (Ainley et al. 2016).

Conclusion

Cardiac interoception deals with the conscious attentional orientation towards cardiac signals. This capacity may have emerged to optimize homeostasis, since it can allow the organism to adapt its behavior (Damasio & Carvalho 2013). However, it is most often the case that we are not paying attention to heartbeats, even though cardiac signals are constantly reaching our brain. In our view, this other perspective on cardiac signals represents the most important contribution to the understanding of how ascending cardiac signals can be processed in the brain and influence cognition.

E. A major role of the insula?

As we have seen, the insula is one of the direct cortical targets of visceral information (Craig 2002, Critchley & Harrison 2013, Saper 2002). In monkeys, cats and rats, vagal inputs reach a region corresponding to the human insular cortex (Saper 2002). The insular cortex of the rat is viscerotopically organized, with cardiopulmonary inputs being located in the posterior (granular) insula (Cechetto & Saper 1987). In humans, visceral inputs reach the mid/posterior insula (Craig 2002, Damasio & Carvalho 2013), but to our knowledge no viscerotopical organization has been demonstrated. It has been hypothesized that visceral signals are then conveyed to anterior insula regions (Craig 2009, Damasio & Carvalho 2013), where activity correlates with cardiac awareness and gray matter volume with a general bodily awareness (Critchley et al. 2004). Interoceptive signals would be integrated with exteroceptive (Farb et al. 2012, Simmons et al. 2013), vestibular (Mazzola et al. 2014) and gustatory (Avery et al. 2015) signals at different stages along this gradient, leading to a general bodily representation in the anterior insula. A functional connectivity study recently showed that in good heartbeat perceivers the connectivity between the right posterior and the right anterior insula during the heartbeat counting task was decreased (Kuehn et al. 2016). This result seems counter-intuitive, but the authors argue that this could reflect a mechanism of noise reduction, by decreasing the processing of other signals (interoceptive, tactile, vestibular…).

Craig developed an influential theory according to which visceral signals are integrated with environmental, hedonic, motivational, social and cognitive signals, in a gradient along the insular cortex giving rise to awareness in the anterior insula (Craig 2009, Strigo & Craig 2016). A recent meta-analysis indeed shows that the same regions of the right anterior insula are active during interoception, emotions and social cognition tasks (Adolfi et al. 2017). However, other regions also appeared to be important, thereby forming a large insular-frontotemporal network. Another meta-analysis has shown the involvement of *both* the anterior and the posterior insula in the processing of cognitive, affective and somatosensory tasks (Beissner et al. 2013). Additionally, even if insular activity is sometimes reported in self-processing studies (Hu et al. 2016, Qin & Northoff 2011, Svoboda et al. 2006), the most consistent structures involved are in the DN (Northoff et al. 2006, Qin & Northoff 2011). Further, as underlined by Damasio and Carvalho, the claim that "the anterior insular

cortex engenders human awareness" is problematic (Damasio & Carvalho 2013). If the anterior insula were *the* neural correlate of awareness, one would expect a bilateral insula lesion to be dramatically detrimental, with consequences possibly similar to vegetative or minimally conscious states. Yet, a patient with a bilateral insula lesion maintained intact feelings (Damasio et al. 2012) while another was able to perform the heartbeat detection task (Khalsa et al. 2009), suggesting at least some degree of preserved awareness. Finally, some studies failed to find modulations of insular activity with some physiological variations (in particular hypoglycemia (Teves et al. 2004)), which contradicts the idea that the insula is *the* neural structure responsible for an integrated representation of the organism. As Craig himself notes (Craig 2009), the anterior insula is very often activated in conjunction with other structures, in particular the anterior cingulate cortex, which could also be important for the integration of visceral signals.

Medford and Critchley argue that the anterior insula should be considered jointly with anterior cingulate activity (Medford & Critchley 2010). In their model, both structures form a functional system. While the anterior insula integrates sensory information (input function), the anterior cingulate cortex re-represents this information and leads to the selection and preparation of appropriate responses (output function). Importantly, the back-projections from the anterior cingulate cortex to the insula would allow this structure to control the integrative function of the anterior insula.

F. Visceral signals and bodily awareness

We have previously discussed aspects of the somatosensory/motor body that contribute to bodily self-consciousness. Somatosensory/motor and visceral signals are conveyed to the brain through similar pathways and both target cortical areas where these signals can be integrated. Now let us address the possible interactions between visceral and somatosensory/motor signals in the context of bodily self-consciousness.

If we consider cardiac signals, no variations in heart rate were observed in relation to agency (David et al. 2011). Similarly, no link has been found between full-

body illusions and cardiac interbeat interval or heart rate variability (Park et al. 2016).

If cardiac activity does not appear to be modulated by these illusions, these illusions, in turn, could be modulated by ascending cardiac signals. Self-identification and self-location measures were increased in the full-body illusion when the virtual body flashed in synchrony with the participant's heartbeats (Aspell et al. 2013). An enhancement of the effect was also found in the rubber hand (Suzuki et al. 2013) and in the enfacement illusion (Sel et al. 2016). HER amplitude can be modulated during experiments on full-body illusions (Park et al. 2016) and enfacement (Sel et al. 2016), depending on the intensity of the illusion. If the flash is in turn synchronous with breathing, participants perceive the location of their breathing as displaced toward the virtual body (Adler et al. 2014). Furthermore, interoceptive abilities modulate the strength of the rubber hand (Suzuki et al. 2013, Tsakiris et al. 2011) and enfacement illusion (Sel et al. 2016), showing that the malleability of the somatosensory/motor body relies on interoceptive signals.

In terms of neuroimaging results, many studies have shown the involvement of the insula in the rubber hand illusion (Tsakiris et al. 2007a), first-person perspective processes (Vogeley et al. 2004) and agency (Farrer & Frith 2002). Moreover, changes in self-location and first-person perspective during the full-body illusion induce changes in the functional connectivity between the left TPJ and the insula (Ionta et al. 2014). Some authors speculate that these insular activations reflect the integration between external and internal signals processed in this region (Blanke et al. 2015, Suzuki et al. 2013, Tsakiris et al. 2011). The most direct evidence for the role of the insula comes from the case of a patient after insular resection (Ronchi et al. 2015), who experienced stronger full-body illusion and had decreased heartbeat awareness. However, this prominent role of the insula was not corroborated in a recent study (Park et al. 2016), which found that HER amplitude differed during the full-body illusion, but only in the posterior cingulate cortex (actually mid-cingulate cortex and supplementary motor area according to the AAL atlas (Tzourio-Mazoyer et al. 2002)).

Conclusion and main questions

The main goal of this thesis was to study the relationship between heart-brain coupling and selfhood. In practice, we looked at heartbeat-evoked responses, as an index of the heart-brain coupling. We hypothesized that this mechanism could take place in different brain regions, in order to tag brain processes as being self-related (articles I (Babo-Rebelo et al. 2016a), II (Babo-Rebelo et al. 2016b) and III in prep.). The default-network and the insula are good candidates for the implementation of this mechanism, but it is interesting to consider the relative importance of these brain structures (article II, (Babo-Rebelo et al. 2016b)).

We wanted to study selfhood beyond explicit bodily self-consciousness, to examine the role of implicit bodily signals in higher-level forms of the self. Our secondary aim was therefore to show how the heart-brain coupling could differentiate the "I" and the "Me", two fundamental self-dimensions that can be expressed in spontaneous thoughts (article I and II, (Babo-Rebelo et al. 2016a,b)). We then tested how this mechanism could translate into a task where thoughts are oriented and where we contrast imagining the self and imagining the other (article III, in prep.).

V. Article I: Neural responses to heartbeats in the default-network encode the self in spontaneous thoughts

A. Technical remarks on heartbeat-evoked responses

1. Confounding artefacts: cardiac-field and pulse artefacts

As we have said, the neural monitoring of the heart can be studied by measuring heartbeat-evoked responses (HERs, Figure 8). These are believed to be of neural origin, i.e. *brain* responses to heartbeats. However, there are two other cardiac-related signals, which are not of neural origin, that are also obtained by locking EEG, iEEG or MEG data to heartbeats. These have to be controlled for or corrected if one wants to properly study HERs.

The first one of these signals is the cardiac-field artefact. The cardiac-field artefact results from the myocardial contraction that can be directly picked up by EEG or MEG sensors. This signal, of cardiac origin, overlaps with the recording of brain activity and is often more prominent in lateral sensors. This artefact appears during the QRS complex phase of the cardiac cycle, during heart contraction, but also during the T-wave, which occurs around 250ms after the R-peak and corresponds to the onset of diastole (Dirlich et al. 1998). Kern and colleagues (Kern et al. 2013) demonstrated that the cardiac-field artefacts in EEG are four times higher in amplitude than cardiac-field artefacts in iEEG. In MEG, the cardiac-field artefact is estimated to have on average 130 ft/cm in amplitude (Jousmäki & Hari 1996).

The second type of artefact is the pulse-related artefact. This artefact is created by pulsating blood vessels, the pulsatile circulation of the cerebrospinal fluid and the resulting pulsatile motion of brain tissue, which can displace recording electrodes and cause impedance changes (Kern et al. 2013). This artefact appears in the form of a slow frequency sinewave or a saw-tooth pattern, which peaks at around 200 ms in non-invasive EEG, but has a more variable timing in iEEG. This artefact was estimated to be six times larger in amplitude in iEEG than EEG data, and did not depend on the distance between the iEEG recording site and the closest blood vessel. In MEG, this artefact is considered to be negligible (Jousmäki & Hari 1996).

Even though some studies on HERs do not discuss the implications of these artefacts (Couto et al. 2015, MacKinnon et al. 2013, Park et al. 2015), it is particularly

important to take them into account. If a task alters the amplitude of the electrocardiogram (ECG) signals or blood pressure it is difficult to disentangle the potential neural effects from the heart-related effects. Indeed, the amplitude of the T-wave and blood pressure may vary during tasks (Gray et al. 2007).

2. *Correction and control of artefacts*

Although correcting or controlling the pulse-related artefact is difficult, the cardiac-field artefact can be very attenuated in different ways.

One way to control for this artefact is to restrict the time window of analysis to the time between two heartbeats, i.e. to the time between the T-wave and the following R-peak (Dirlich et al. 1998, Gray et al. 2007, Immanuel et al. 2014, Leopold & Schandry 2001, Park et al. 2014, Schulz et al. 2013b, 2015a,b). This time window is free of the cardiac-field artefacts (Dirlich et al. 1997, 1998). HERs can then be computed locked to the peak of the T-wave of the ECG, since it is closer to the time window of interest (Park et al. 2014). Neural responses to heartbeats could potentially happen earlier as well, but it would be difficult to be sure they are from neural origin.

In addition, cardiac-related artefacts can be corrected using independent component analysis (Canales-Johnson et al. 2015, Luft & Bhattacharya 2015, Terhaar et al. 2012). This technique decomposes the brain signal into different components and computes the coherence between each component and the simultaneous ECG signal. Components having a high coherence with the ECG signal can then be removed from the data. However, this technique may be harsh in the sense that it may also remove neural signals which appear locked to heartbeats. Some authors have directly subtracted the ECG signal to EEG, to control for the cardiac-field artefact (Couto et al. 2014, Fukushima et al. 2011, Montoya et al. 1993, Schandry & Montoya 1996), but this method may not sufficiently take into account that the influence of the cardiac-field artefact may differ across the sensors. The Hjorth source derivation method, which subtracts from each electrode the weighted activity of the surrounding electrodes, has also been applied to attenuate the cardiac-field artefact (Montoya et al. 1993, Pollatos & Schandry 2004, Pollatos et al. 2005a, Shao et al. 2011), but may also be insensitive to subtle changes in the artefact across sensors.

In order to account for the cardiac-field artefact, one can estimate the contribution of heart electrical activity by analyzing ECG activity directly (Gray et al. 2007, Lechinger et al. 2015, Park et al. 2014). As ECG recordings differ depending on the placement of the acquisition electrodes, some authors have recorded ECGs from several electrodes placed around the neck.

Baseline corrections are usual in evoked responses analyses, but in HER analysis, although sometimes applied (Montoya et al. 1993), they are debatable, since they are taken either during the time window of the cardiac-field artefact or during the preceding heartbeat, i.e. during the preceding HER.

B. Abstract in French

Le réseau du mode par défaut (*default-network*, DN) a été associé au soi, mais aussi au suivi de l'état corporel par le cerveau lors de la régulation des fonctions autonomes. Nous avons émis l'hypothèse que ces deux fonctions du DN, apparemment distinctes, pourraient être couplées, en accord avec les théories proposant que le soi est ancré dans le suivi cérébral des organes internes, comme le cœur. Nous avons mesuré en magnétoencéphalographie les réponses cérébrales évoquées par battements cardiaques, pendant que les participants laissaient libre cours à leurs pensées. A des moments aléatoires, un stimulus visuel apparaissait à l'écran. Les participants devaient alors évaluer selon quatre échelles la pensée qui venait d'être interrompue. Ils devaient évaluer dans quelle mesure ils étaient engagés en tant que le sujet de cette pensée (le « Je ») et dans quelle mesure cette pensée était introspective (se référait à « Moi »). Nous avons observé que l'amplitude des réponses évoquées aux battements cardiaques dans le réseau du mode par défaut variait selon que les pensées étaient plus ou moins en rapport avec le soi. La dimension « Je » était associée au cortex cingulaire postérieur / précuneus ventral, et la dimension « Moi » au cortex préfrontal ventromédian. De plus, nous n'avons pas observé de variations dans les mesures physiologiques (rythme cardiaque, variabilité du rythme cardiaque, diamètre pupillaire, activité électrodermale, rythme et phase respiratoire) ni dans la puissance du rythme alpha. Nos résultats démontrent un lien direct entre le soi et les réponses aux battements cardiaques dans le réseau du mode par défaut et donc soutiennent les théories qui ancrent le soi au suivi des signaux viscéraux par le cerveau.

C. Article

Behavioral/Cognitive

Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts

Mariana Babo-Rebelo,[1] Craig G. Richter,[1,2] and Catherine Tallon-Baudry[1,3]

[1]Laboratoire de Neurosciences Cognitives (ENS-INSERM), Département d'Etudes Cognitives, Ecole Normale Supérieure-PSL Research University, 75005 Paris, France, [2]Ernst Strüngmann Institute for Neuroscience in Cooperation with Max Planck Society, 60528 Frankfurt, Germany, and [3]Centre de NeuroImagerie de Recherche CENIR, Université Pierre et Marie Curie-Paris 6 UMR-S975, Inserm U975, CNRS UMR 7225, Groupe Hospitalier Pitié-Salpêtrière, 75013 Paris, France

The default network (DN) has been consistently associated with self-related cognition, but also to bodily state monitoring and autonomic regulation. We hypothesized that these two seemingly disparate functional roles of the DN are functionally coupled, in line with theories proposing that selfhood is grounded in the neural monitoring of internal organs, such as the heart. We measured with magnetoencephalograhy neural responses evoked by heartbeats while human participants freely mind-wandered. When interrupted by a visual stimulus at random intervals, participants scored the self-relatedness of the interrupted thought. They evaluated their involvement as the first-person perspective subject or agent in the thought ("I"), and on another scale to what degree they were thinking about themselves ("Me"). During the interrupted thought, neural responses to heartbeats in two regions of the DN, the ventral precuneus and the ventromedial prefrontal cortex, covaried, respectively, with the "I" and the "Me" dimensions of the self, even at the single-trial level. No covariation between self-relatedness and peripheral autonomic measures (heart rate, heart rate variability, pupil diameter, electrodermal activity, respiration rate, and phase) or alpha power was observed. Our results reveal a direct link between selfhood and neural responses to heartbeats in the DN and thus directly support theories grounding selfhood in the neural monitoring of visceral inputs. More generally, the tight functional coupling between self-related processing and cardiac monitoring observed here implies that, even in the absence of measured changes in peripheral bodily measures, physiological and cognitive functions have to be considered jointly in the DN.

Key words: default network; heartbeat-evoked responses; MEG; self; spontaneous cognition

Significance Statement

The default network (DN) has been consistently associated with self-processing but also with autonomic regulation. We hypothesized that these two functions could be functionally coupled in the DN, inspired by theories according to which selfhood is grounded in the neural monitoring of internal organs. Using magnetoencephalography, we show that heartbeat-evoked responses (HERs) in the DN covary with the self-relatedness of ongoing spontaneous thoughts. HER amplitude in the ventral precuneus covaried with the "I" self-dimension, whereas HER amplitude in the ventromedial prefrontal cortex encoded the "Me" self-dimension. Our experimental results directly support theories rooting selfhood in the neural monitoring of internal organs. We propose a novel functional framework for the DN, where self-processing is coupled with physiological monitoring.

Introduction

The default network (DN) has been consistently associated with self-related processing in fMRI studies (Buckner et al., 2008; Qin and Northoff, 2011; Andrews-Hanna et al., 2014). However, the DN is also involved in central autonomic processing (Thayer et al., 2012; Beissner et al., 2013) and includes prefrontal visceral cingulate areas (Vogt and Derbyshire, 2009), which respond to heartbeats (Park et al., 2014) and modulate heart rate when stimulated (Van Eden and Buijs, 2000). This colocalization between a

Received Jan. 25, 2016; revised April 14, 2016; accepted May 19, 2016.

Author contributions: M.B.-R. and C.T.-B. designed research; M.B.-R., C.G.R., and C.T.-B. performed research; M.B.-R., C.G.R., and C.T.-B. analyzed data; M.B.-R. and C.T.-B. wrote the paper.

This work was supported by European Research Council ANR-BLAN-12-BSH2-0002-01 under the European Union's Horizon 2020 research and innovation program Grant agreement 670325 to C.T.-B., and ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL. M.B.-R. was supported by Fundação para a Ciência e Tecnologia Grant SFRH/BD/85127/2012. We thank Christophe Gitton for technical assistance during MEG recordings and Antoine Ducorps for help with MEG data preprocessing.

The authors declare no competing financial interests.

This article is freely available online through the *J Neurosci* Author Open Choice option.

Correspondence should be addressed to Dr. Catherine Tallon-Baudry, Laboratoire de Neurosciences Cognitives ENS-INSERM, 29 Rue d'Ulm, 75005 Paris, France. E-mail: catherine.tallon-baudry@ens.fr.

DOI:10.1523/JNEUROSCI.0262-16.2016

cognitive role in selfhood and a physiological role in autonomic function, although sometimes noted (Buckner et al., 2008; Iacoboni and Hasson, 2011), remains largely unexplained. Several theories propose that the neural monitoring of visceral signals participates in the experience of selfhood by contributing to an integrated neural representation of the organism as a unified entity (e.g., as a self) (Damasio, 1999; Craig, 2009; Park and Tallon-Baudry, 2014). We thus hypothesized that cardiac monitoring and self-processing are functionally coupled in the DN.

We measured heartbeat-evoked responses (HERs) using magnetoencephalography (MEG) in a thought sampling paradigm (Hurlburt and Heavey, 2001), where participants rated the self-relatedness of spontaneous thoughts. HERs (Schandry et al., 1986) are obtained by averaging electrophysiological data locked to heartbeats (Schandry and Montoya, 1996; Gray et al., 2007; Kern et al., 2013; Park et al., 2014; Lechinger et al., 2015). At each heartbeat, information about heart contraction is transmitted, through vagal and spinal pathways, to the neocortex (Armour and Ardell, 2004; Critchley and Harrison, 2013), where it elicits transient HERs, in the right insula and somatosensory cortices (Pollatos et al., 2005; Kern et al., 2013; Canales-Johnson et al., 2015), but also in the DN (Park et al., 2014). The participants' task was to fixate a point on a screen and to let their thoughts develop freely until the appearance of a visual stimulus (see Fig. 1*B*). Participants then rated the self-relatedness of the interrupted thought, as detailed in the next paragraph, its emotional intensity, as well as whether the thought related to past, present, or future events (D'Argembeau et al., 2008; Tusche et al., 2014; Couto et al., 2015). Our objective was to test whether the amplitude of HERs during the thought systematically covaried with its self-relatedness (see Fig. 1*C*), and whether this mechanism engaged the DN.

Self-relatedness in spontaneous thoughts can be expressed as the "I" (i.e., the agent or subject in the thought) or as the "Me" (i.e., when participants think about themselves) (see Fig. 1*A*). The "I" scale described participants' engagement as the subject of the thought (i.e., acting, feeling, or perceiving) from the first-person perspective. "I" ratings were high for thoughts, such as "I have to make a phone call" or "I am thirsty," and low for thoughts, such as "It's raining" or "He is coming tomorrow." The "Me" scale described the content of the thought. Ratings were high when participants were thinking about themselves, as in "I am thirsty" or "I should be more concerned," and low when the thought was directed toward something or someone else, as in "It's raining" or "He is coming tomorrow." The conceptual distinction between these two self-dimensions (James, 1890; Legrand and Ruby, 2009; Christoff et al., 2011; Gallagher, 2012) has been emphasized and might prove experimentally useful (Powell et al., 2010; Christoff et al., 2011).

Materials and Methods

Participants. 20 right-handed volunteers participated in this study after giving written informed consent and were paid for their participation. The study was approved by the ethics committee CPP Ile de France III. Four participants were excluded from analysis: excessive eye-movement ($n = 2$), excessive body movements ($n = 1$), extremely fast heart rate ($n = 1$) (mean interbeat interval of 687 ms, >2 SDs faster than the average interbeat interval in the other participants). Sixteen participants were thus included in the analysis (8 male; mean age: 24.1 ± 0.6 years).

Thought-sampling task. Each trial of the thought-sampling task consisted of a fixation period (central black dot, radius 0.13° of visual angle, surrounded by a black circle, radius 0.38° of visual angle, on a gray background) followed by a visual stimulus (8 white dots centered on fixation, radius 0.13° of visual angle, arranged in a square of 1.54° of visual angle, presented for 200 ms). Fixations ranged from 13.5 to 30 s in 1.1 s steps and were randomized in each block, so that participants could not guess when they would be interrupted. Participants were asked to let their mind wander as naturally as possible during fixation while avoiding structured thinking (e.g., singing, counting…), and to press a button in response to the visual stimulus. Then, they rated the thought they were having at the moment of the stimulus display, along four continuous scales. The "Actor/Author" and "Content" scales targeted the "I" and "Me" dimensions of the self, respectively. The "Time" scale was used to report whether the thought referred to past, present, or future events, whereas the "Valence" scale was used to determine whether the thought was pleasant or unpleasant (for precise instructions on the meaning of the four scales, see Training procedure and instructions). Participants responded by moving a cursor to the left or to the right of the continuous scales (range: 1–202, 1.5 steps, pressing left and right buttons with their right index or middle finger, respectively) and validated their choice with the right thumb, within 20 s per scale. The order of the scales was constant for a given participant but randomized between participants. Participants could skip the ratings if they did not have any clear thought when the stimulus appeared or if they did not know how to rate the thought. If a trial was skipped, a new one was added to the block, unbeknownst to the participant.

Training procedure and instructions. In preparation of the MEG thought-sampling task, 20 pilot participants performed the thought-sampling task but, in addition to ratings, had to verbally report the content of their spontaneous thoughts at the end of each trial. We selected 32 descriptions of thoughts from this pilot study, to train and test the group of participants used for the MEG experiment.

To make sure that MEG participants understood the task, each participant visited the laboratory a few days before the MEG session and was instructed about the task and trained on the scales by rating 22 of the 32 descriptions of thoughts obtained in the pilot study. If necessary, ratings were discussed with the experimenter to clarify the meaning of the scales. We detail below the rating instructions and provide a number of examples.

The "Actor/Author" scale targeted the "I" dimension of the self ("I" scale) and evaluated the degree to which the participant was seeing or feeling himself/herself as the actor or author during the thought. Participants were instructed to use high ratings ("+") when they were adopting their own perspective (i.e., when they were the protagonist or the agent of thought), as in "Tonight I'm doing the laundry." Low ratings ("−") were used when someone else was the protagonist of the thought ("His office is far away") or nobody in particular ("It's raining"). Participants were asked to use the whole extent of the scale, including intermediate levels, to better characterize their degree of involvement as the "I" during the thought.

The "Content" scale targeted the "Me" dimension of the self ("Me" scale) (i.e., how much the thought was focused on the participant himself/herself or on something external). The "Me" extreme of the scale was to be used when participants were thinking about themselves, about their feelings, body, or mood, as in "I'm hungry," "I should be more concerned," or "I'm bored." The "External" extreme was to be used when participants were thinking about something that was external to them, as for instance "It's raining" or "What was the title of the book that Peter recommended?"

Critically, thoughts where participants were the protagonist but were not focusing on themselves had to be rated high in the "I" scale and low on the "Me" scale. This would be the case for "I'll go to the bakery because there is no more bread at home," where I am the protagonist but I am not focusing on my feelings. Ratings are different if the thought is "I'll go to the bakery because I'm craving for a croissant." In this example, I am again the protagonist but I am this time focusing on myself, specifically on my desire for a croissant. A high rating should thus be used in both scales. Conversely, thoughts where participants were thinking about the opinion someone else had about them were to be rated high on the "Me" scale and low on the "I" scale (e.g., "He likes me").

The "Time" scale was used to report whether thoughts referred to a past, present, or future event. Participants rated events that occurred a few weeks ago on the lower 20% of the scale, a few days ago between 20% and 40%, present and a few hours before/after between 40% and 60%, in a few days between 60% and 80%, and in a few weeks >80%.

The "Valence" scale was used to report whether thoughts were pleasant ("positive"), neutral (center of the scale) or unpleasant ("negative"). Participants were instructed to try and finely evaluate their thoughts by using all degrees of the scale. They were asked to use the higher and lower end of the scale for everyday life situations strongly positive or negative, not the most positive or most negative thought they ever had.

After reading and discussing the instructions and rating examples, participants performed 6 trials of the thought-sampling task to familiarize themselves with the procedures, and could further clarify the scales with the experimenter if necessary.

Experimental procedure. Just before the MEG recording, participants were reminded of the instructions and asked to rate 10 new example thoughts. Ratings were discussed with the experimenter to ensure task comprehension. Participants then performed a practice block of 6 trials of the thought-sampling task, followed by 5 blocks of 16 trials during which MEG and physiological data were acquired. This was followed by a 12 min resting-state sequence, where participants maintained fixation while avoiding structured thinking. After MEG recordings, participants were tested on their interoceptive abilities by counting their heartbeats (Schandry, 1981) while focusing on their bodily feelings and fixating on the screen, in six blocks of different durations (30, 45, 60, 80, 100, 120 s, order randomized between participants), without feedback on performance. Participants then completed a questionnaire about the experiment as well as a French version of the Daydreaming Frequency Scale (Giambra, 1993), and the Self-Consciousness Scale (Fenigstein et al., 1975). Eighteen months later, participants completed the Trait Anxiety Inventory (Spielberger et al., 1983).

Recordings. Continuous MEG data were acquired using a whole-head MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX, sampling rate of 1000 Hz, online low-pass filtered at 330 Hz). ECG data (0.03–330 Hz) were obtained from 7 electrodes placed around the base of the neck and referenced to a left abdominal location to estimate the cardiac field artifact as best as possible. The ground electrode was located on the left costal margin. Two ECG electrodes were placed over the left and right clavicles, two over the top of the left and right shoulders, two over the left and right supraspinatus muscle, and one over the upper part of the sternum. Interbeat intervals consisted of the average time distance between the two T peaks preceding the visual stimulus and the heart rate variability corresponded to the SD of the interbeat intervals. Electrodermal activity was recorded via two electrodes on the sole of the left foot, and respiratory activity was recorded via a respiratory belt positioned around the chest, at the level of armpits (respiratory transducer TSD201 BIOPAC system; removed for the heartbeat counting task). Both signals were low-pass filtered at 330 Hz. Horizontal and vertical eye position and pupil diameter were monitored using an eye-tracker (EyeLink 1000, SR Research) and recorded simultaneously with MEG, ECG, electrodermal activity, and respiratory data. Stimuli were presented on a semitranslucent screen at an 85 cm viewing distance.

MEG data preprocessing. Continuous MEG data were denoised using temporal signal space separation (as implemented in MaxFilter) and bandpass filtered between 0.5 and 40 Hz (fourth-order Butterworth filter). Blinks and saccades >2 degrees were identified by the Eyelink system. Epochs contaminated by large movement or muscle artifacts were visually detected. Independent component analysis (ICA), as implemented in the Fieldtrip toolbox (Oostenveld et al., 2011), was used to correct for the cardiac field artifact, for both magnetometer and gradiometer signals, based on epochs of −1.5 to 1.5 s around the T peaks of interest that were devoid of movement, muscle, blink, or saccade artifacts. Because temporal signal space separation induces rank deficiency, we defined the number of ICA components by first computing a principal component analysis (PCA). We then removed all the independent components with a mean pairwise phase consistency (Vinck et al., 2010) with the lead II ECG signal, in the 0–25 Hz range, >0.2 (from 0 to 2 components per participant). ICA-corrected MEG data were then low-pass filtered at 25 Hz (fourth-order Butterworth filter).

HERs. We first detected the R peaks by correlating the ECG with a template QRS complex defined on a subject-by-subject basis and identifying the local maximum within the episodes of correlation >0.7. T peaks were then detected by first correlating the ECG with a template of the T peak, followed by identifying the local maxima within episodes with a correlation >0.5 (except for one subject: 0.3) that followed an R peak by at most 0.4 s. R and T peak detection was visually verified in all subjects. The two T peaks preceding the visual stimulus by at least 400 ms were used for HER computation.

By taking two heartbeats per trial, we increased the signal-to-noise ratio while assuming a realistic duration for a stable thought (mean duration between the last-but-one heartbeat and the visual stimulus: 1.80 ± 0.032 s). We rejected epochs (from 0.2 s before to 0.5 s after the selected T peaks) contaminated with saccades >2° of visual angle from fixation, blinks and movement, or muscular artifacts.

Trial classification. We used a median split to label trials as "high" or "low" on each scale. Only trials with at least one artifact-free HER were considered in the median split. If ratings were equal to the median, they were arbitrarily assigned to the "high" group, a procedure that resulted in marginally different trial numbers in the "high" and "low" groups (mean difference in number of trials: "I" scale = 1.8 ± 0.5%, "Me" scale = 1.2 ± 0.4%, Time = 8.0 ± 2.0%, Valence = 5.1 ± 1.3%). Artifact-free HERs corresponding to "high" and "low" ratings were computed by averaging magnetometer data across heartbeats, from 0.1 s before the T peak to 0.4 s after the T peak.

Cluster-based permutation procedure. The significance of the difference in HERs between "high" and "low" ratings on the four scales was tested on magnetometer signals, in the artifact-free time window 80–350 ms after the T peak, using a cluster-based permutation *t* test (Maris and Oostenveld, 2007). This method does not require the definition of any a priori spatial or temporal regions and intrinsically corrects for multiple comparisons in time and space. For each scale, a *t* value is computed between HERs for "high" and "low" ratings. Individual samples with a *t* value corresponding to a *p* value below a selected threshold ($p < 0.01$, two tailed) are clustered together based on temporal and spatial adjacency. The cluster is characterized by the sum of the *t* values of the individual samples. To establish the likelihood that a cluster was obtained by chance while controlling for the fact that four different scales were tested, we shuffled the "high" and "low" labels 10,000 times and repeated the clustering procedure on each scale selecting the maximum positive cluster-level statistic and the minimum negative cluster-level statistic across the four tests. For each scale, the Monte Carlo *p* value corresponds to the proportion of elements in the distribution of maximal (or minimal) cluster-level statistics that exceeds (or is inferior to) the originally observed cluster-level test statistics. Cluster amplitude corresponds to the average of the magnetometer data across significant sensors in the significant time window. This procedure was also applied at the source level, independently on the two self-related scales, on currents averaged over the time windows identified by the sensor level test, on the 15,002 vertices of the cortical surface model. The same clustering procedure, with the same thresholds, was also applied on ECG data, separately on vertical and horizontal derivations.

PCA of the "I" and "Me" ratings. The "I" and "Me" ratings were rank-based inverse normal transformed (Bishara and Hittner, 2012) and *z*-scored for each participant. To determine the dimension capturing the variance common to both scales, PCA was performed using the MATLAB function princomp (The MathWorks). The scores of each participant were projected on the first PCA component and labeled as "high" or "low" relative to the median of this general self-relatedness scale.

General linear model (GLM). GLMs were applied to the magnetometer data in the "I" and "Me" clusters, with the rank-based inverse normal transformed (Bishara and Hittner, 2012) and *z*-scored ratings on the four scales as regressors. Each regressor was Gram-Schmidt orthogonalized with respect to the preceding regressors specified in the model. Shared variance between regressors r_i and r_{i+1} is hence assigned to r_i, whereas r_{i+1} retains only its unique variance. We computed two GLMs, each with a different order of regressors: model 1 with the regressor order "Me"-"I"-Time-Valence ratings to test whether the unique variance of the "I"

ratings accounts for the "I" cluster; model 2 with the regressor order "I"-"Me"-Time-Valence ratings to test whether the unique variance of the "Me" ratings accounts for the "Me" cluster. For each model, β values of each regressor were averaged over the channels and time window of the significant cluster being tested. The crucial test was whether the unique variance of the second regressor accounted for the data, after the shared variance with the first regressor has been removed and assigned to the first regressor by the orthogonalization procedure. This test was achieved by testing whether the β corresponding to the second regressor significantly differed from 0 across participants.

To assess the degree of collinearity between the four regressors, we additionally computed variance inflation factors, for each subject, between each scale and the other three scales.

Evidence in favor of an absence of differences. Bayes factors were computed to evaluate evidence in favor of the null hypothesis. For paired *t* tests, we computed the maximum log-likelihood of the model in favor of the "null" hypothesis and the model in favor of the "effect" hypothesis. The group-level random-effect Bayes factor was computed with the prior reference effect corresponding to an effect differing from 0 under a *t* test with a *p* value of 0.05. We then used the Bayesian information criterion to compare the two models and compute the corresponding Bayes factor. We also computed Bayes factors on the regression between personality factors and our results by using the online Bayes factor calculator tool (http://pcl.missouri.edu/bayesfactor), which is based on Liang et al. (2008).

As a rule of thumb, a Bayes factor >3.2 provides substantial evidence in favor of the null hypothesis, whereas a Bayes factor <3.2 does not provide enough evidence for or against the null hypothesis (Kass and Raftery, 1995).

Surrogate heartbeats. To demonstrate that the observed effects were locked to heartbeats, we checked whether the differences between "high" and "low" trials could be obtained with the same sampling of the neural data but unsynchronized with heartbeats. We created 100 permutations of heartbeats, where the timing of the pair of heartbeats of trial *i* in the original data was randomly assigned to trial *j*. The same criteria for rejecting artifactual epochs, median-splitting of the data according to behavior and computing of HERs were applied. For each permutation, we obtained a set of neural responses to surrogate heartbeats and computed the cluster summed *t* statistics as described above. For each permutation, we extracted the largest positive sum of *t* values in the comparison between "high" and "low" ratings on the "Me" scale, and the smallest negative sum of *t* values for the "I" scale, and compared the distribution of those surrogate values with the observed original sum of *t* values.

Anatomical MR acquisition and preprocessing. An anatomical T1 scan was acquired for each participant, on a Siemens TRIO 3T ($n = 13$) or Siemens VERIO 3T ($n = 3$) scanner. Segmentation of the data was processed with automated algorithms provided in the FreeSurfer software package (Fischl et al., 2004) (http://surfer.nmr.mgh.harvard.edu/). Segmentations were visually inspected and edited when necessary. The white-matter boundary was determined using FreeSurfer and was used for subsequent minimum-norm estimation.

Source reconstruction and comparison with fMRI findings. Source localization and surface visualization were performed with the BrainStorm toolbox (Tadel et al., 2011). After coregistration between the individual anatomy and MEG sensors, cortical currents were estimated using a distributed model consisting of 15,002 current dipoles from the combined time series of magnetometer and gradiometer signals using a linear inverse estimator (weighted minimum-norm current estimate, signal-to-noise ratio of 3, whitening PCA, depth weighting of 0.5) in a single-sphere head model. Dipole orientations were constrained to the individual MRIs. Cortical currents were then averaged over the time windows for which a significant difference between "high" and "low" responses on the "I" and "Me" scales was identified in sensor space, spatially smoothed (FWHM 7 mm), and projected to a standard brain model (Colin27, 15,002 vertices). Reliable differences in dipole current values were identified using the same cluster-based procedure as described for the sensor level analysis applied to the 15,002 vertices.

The coordinates of the vertex corresponding to the maximal *t* value in the cluster were reported. Anatomical descriptions are based on the Tzourio-Mazoyer parcellation (Tzourio-Mazoyer et al., 2002). The functional connectivity map was obtained in Neurosynth (Yarkoni et al., 2011) using the coordinates of the "Me" cluster as the seed region (threshold for visualization: Pearson correlation $r = 0.19$). The default network map (Laird et al., 2011) was converted from Talairach to MNI coordinates using the functions Normalize and Image Calculator in SPM8. The final figure was created with Mango (http://ric.uthscsa.edu/mango/).

Physiological and arousal measures processing. In addition to the seven vertical ECG signals recorded, we offline computed the seven bipolar horizontal derivations between adjacent electrodes. ECG measures were preprocessed and analyzed in an identical manner to the MEG data.

Respiratory data were epoched from -12 to 7 s around the visual stimulus. Artifactual epochs were detected visually and excluded from analysis. Epochs were then mean-centered by subtracting the mean value in the 7 s preceding the visual stimulus and 0 crossings were detected. Two successive 0 crossings defined a respiratory cycle. To test whether respiratory phase could impact the differential HERs observed, for each heartbeat of the analysis, we computed the respiratory phase corresponding to 132 and 313 ms after the T peak, which correspond to the center of the significant time windows for "Me" and "I," respectively. We then computed the phase bifurcation index (Busch et al., 2009) separately for each scale, to test for differences in phase distribution between "high" and "low" ratings, for each participant. Finally, we tested whether this measure differed from 0 across participants, which would indicate that heartbeats would be locked to different respiratory phases in trials rated as "high" and in trials rated as "low" in the corresponding scale.

Blinks were automatically detected with the Eyelink software. The time windows identified by the Eyelink system as containing a blink were extended by 80 ms on each side. We further identified and rejected all variations in pupil diameter >200 (arbitrary units) in a 200 ms time window. To analyze pupil diameter, portions of data containing blinks were linearly interpolated and a low-pass fourth-order Butterworth filter at 10 Hz was applied. Data were then epoched from 80 ms after the last-but-one T peak preceding each visual stimulus and 1.3 s after the visual stimulus. Epochs with $>30\%$ noisy data (blinks) were excluded from analysis. The remaining epochs were *z*-scored.

Electrodermal activity was low-pass filtered at 10 Hz (fourth-order Butterworth filter) and the extracted epochs were *z*-scored before averaging.

To compute alpha power, ICA-corrected MEG data were bandpass filtered between 8 and 12 Hz (fourth-order Butterworth filter) and the corresponding alpha-band power was computed using the Hilbert transform. Data from the 15 sensors showing the largest alpha power at the group level were averaged. Epochs containing blinks or muscle artifacts were discarded before averaging.

Pupil diameter, electrodermal activity, and alpha power data were averaged in each epoch from 80 ms after the last-but-one T peak preceding the visual stimulus to 400 ms before the visual stimulus. Then the mean value of each epoch was averaged for "high" trials and "low" trials, along the "I" and "Me" scales.

Results

Task comprehension

Before the MEG experiment, participants were tested for task comprehension by rating a list of 10 written example thoughts. Between-participant rating consistency was high on all four scales (Cronbach's alpha coefficient, "I": 0.9981; "Me": 0.9822; Time: 0.9790; Valence: 0.9882), showing that participants understood the instructions and applied similar criteria when using the scales. In a debriefing questionnaire after the MEG experiment, participants reported that it was easy to mind wander spontaneously, that interrupted thoughts were stable and precise, and importantly, that it was easy to use the scales to rate their own thoughts (Table 1). In addition, participants were given the possibility to skip a trial if they were not sure how to use the scales (Fig. 1*B*). Participants skipped only 2.8 ± 1.1 trials (range across participants from 0 to 15 trials).

Table 1. Debriefing questions on the experiment

Questions	Response scale	Response (mean ± SEM)
Was it difficult to use the scales?	1 very easy, 7 very difficult	2.5 ± 0.2
Did you hesitate a lot when rating your thoughts, or were your rating decisions easy, immediate, natural?	1 immediate decisions, 7 difficult decisions	2.5 ± 0.2
Was it hard to catch the interrupted thought?	1 rarely, 7 frequently	3.1 ± 0.3
Were your thoughts too fast to be caught or were they stable and easily graspable?	1 thoughts were too fast, 7 thoughts were stable and slow	5.2 ± 0.3
Were you able to let your thoughts wander?	1 rarely, 7 always	5.4 ± 0.2
Were your thoughts precise enough so you could rate them?	1 very imprecise, 7 very precise	5.2 ± 0.3

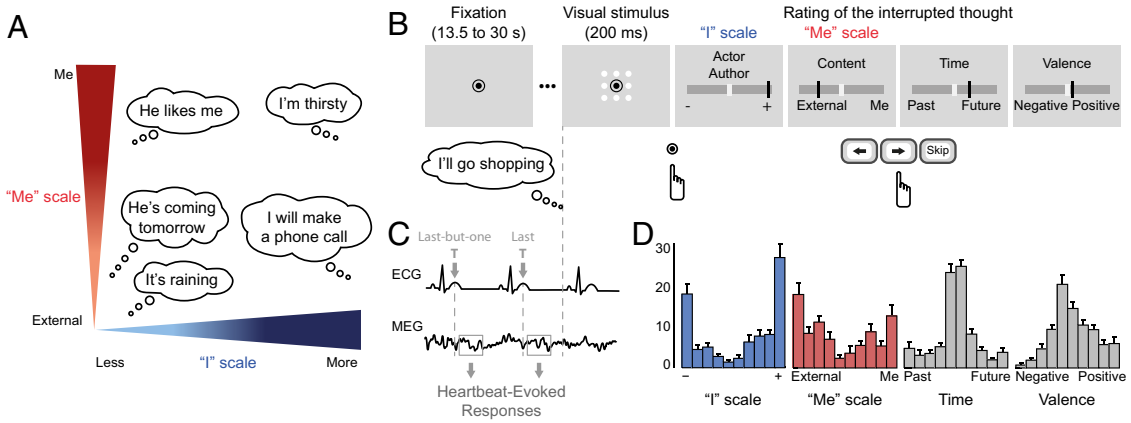


Figure 1. Experimental paradigm and behavior. ***A***, Examples of thoughts along the two scales of self-relatedness. The "Me" scale described the content of the thought oriented either toward oneself or toward an external object, event, or person. The "I" scale described the engagement of the participant as the protagonist or the agent in the thought. ***B***, Time course of a trial. Each trial consisted of a fixation period (13.5–30 s, randomized) interrupted by a visual stimulus. During fixation, participants were asked to let their thoughts develop freely while avoiding structured thinking (e.g., singing, counting…). Participants pressed a button in response to the visual stimulus and had to remember the thought that was interrupted by the visual stimulus. Then, they rated this thought along four scales ("I," "Me," Time, and Valence). Participants could also skip the ratings if the interrupted thought was unclear or if they were not sure how to use the scales. ***C***, Selection of MEG data locked to the two T peaks of the ECG preceding the visual stimulus to compute heartbeat-evoked responses during the thought. ***D***, Distribution of ratings on the scales, across all participants ($n = 16$) and thoughts ($n = 80$ per participant). Error bars indicate SEM.

Behavioral results

The distributions of the ratings of spontaneous thoughts along the four scales are presented in Figure 1*D*. Ratings on the "I" scale, but not on the "Me" scale, were slightly biased toward high self-relatedness ("I" scale median: 135.4 ± 12.4 SEM, *t* test against the middle of the scale; range: 1–202, middle: 101.5: $t_{(15)} = 2.74$, uncorrected $p = 0.015$; "Me" scale: median: 84.7 ± 11.5, $t_{(15)} = -1.46$, uncorrected $p = 0.17$). Time ratings were centered on "present" (median: 101.6 ± 1.2, $t_{(15)} = 0.079$, uncorrected $p = 0.94$), and Valence ratings were slightly biased toward positive contents (median: 110.7 ± 3.2, $t_{(15)} = 2.86$, uncorrected $p = 0.012$).

We then tested for correlations between scales. Across participants, the mean correlation between the ratings on the "I" and "Me" scales was significantly positive (mean Fisher *z*-transformed Pearson $r = 0.85 \pm 0.06$, two-tailed *t* test against 0, $t_{(15)} = 13.90$, Bonferroni corrected for the 6 correlations tested $p = 3 \times 10^{-9}$), as well as the correlation between ratings along the "I" and time scales (mean $r = 0.14 \pm 0.03$, $t_{(15)} = 5.21$, Bonferroni corrected $p = 6 \times 10^{-4}$). None of the other between-scale correlations was significant (mean $|r| < 0.045$, Bonferroni corrected $p = 1$). We created two scales meant to target two different aspects of the self. Given the correlation between the ratings on the two self-related scales, we also considered the alternative hypothesis that the two scales reflect the same underlying unitary notion of the self.

HERs covary with the self-relatedness of spontaneous thoughts

We computed HERs by averaging brain activity locked to the T peak of each of the two heartbeats preceding the visual stimulus, in trials with a "high" rating versus trials with a "low" rating, on each scale (median-split of the behavioral data). Despite the correlation between the two self-related scales, 25.1 ± 1.9% of the trials (corresponding to 19.9 ± 1.5 trials for each subject) were classified differently on the two self-related scales (i.e., "high" on one scale and "low" on the other). For each of the four scales, we compared HERs for "high" and "low" trials in the time window 80–350 ms after the T peak, which is devoid of the cardiac field artifact (Dirlich et al., 1998), using a clustering procedure (Maris and Oostenveld, 2007) that identifies significant differences across sensors and time points while correcting for multiple comparisons.

HERs differed significantly between "high" and "low" trials on the "I" scale (cluster sum(t) = −1173.8, Monte Carlo $p = 0.0313$), 298–327 ms after T peak, over medial posterior sensors (Fig. 2*A*,*B*). Moreover, the cluster amplitude for trials rated as "low" or "high" corresponded to fluctuations around a baseline cluster amplitude reference value obtained during a subsequent 12 min resting-state session (Fig. 2*C*; "high": −22.9 ± 4.7 fT, baseline: 0.6 ± 2.1 fT, "low": 17.9 ± 5.4 fT; paired *t* test, "high" vs baseline: $t_{(15)} = -4.2$, $p = 0.0016$; "low" vs baseline: $t_{(15)} = 3.34$, $p = 0.0088$, Bonferroni corrected for the two tests against baseline).

HERs also differed between "high" and "low" trials on the "Me" scale (sum(t) = 1480.6, Monte Carlo $p = 0.0097$), but over

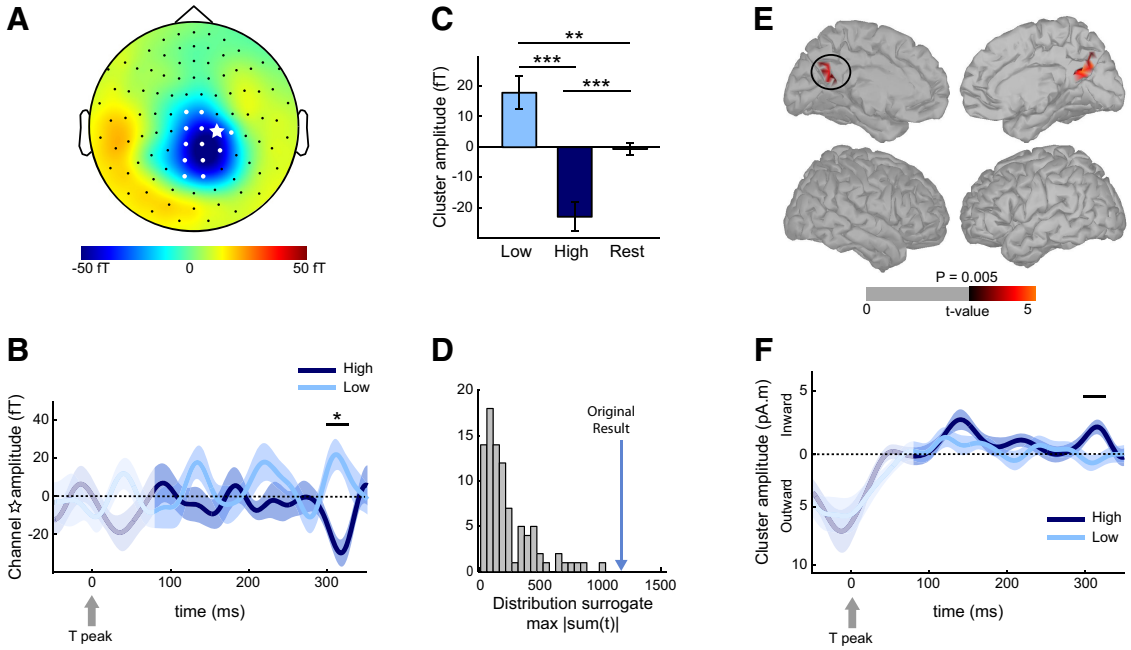


Figure 2. Differential HERs for "high" and "low" ratings on the "I" scale. ***A***, Topographical map of the HER difference between "high" and "low" ratings on the "I" scale, grand-averaged across 16 participants, in the 298–327 ms time window in which a significant difference was observed (Monte Carlo $p = 0.0313$, corrected for multiple comparisons). White dots represent the sensors contributing to the significant cluster. ***B***, Time course of the HER (± SEM) for "high" and "low" ratings on the "I" scale at the sensor indicated in ***A*** (white star). The signal that might be residually contaminated by the cardiac artifact appears in lighter color (before 80 ms, not included in the analysis). Black bar represents the time window in which a significant difference was observed. ***C***, HER cluster amplitude, during thoughts rated as "high" or "low" along the "I" scale, and during a separate eyes-open resting state session. Cluster amplitude during rest was intermediate between cluster amplitude during thoughts rated as "high" ($p = 0.0016$) and cluster amplitude during thoughts rated as "low" ($p = 0.0088$). ***D***, Histogram of the distribution of the maximal cluster *t* statistic (difference between "high" and "low" trials) obtained for the 100 permutations of surrogate heartbeats. The original cluster *t* statistic (arrow) lies outside the distribution of statistics obtained on surrogate data. ***E***, Neural sources of the differential HERs for thoughts rated as "high" or "low" on the "I" scale. Only the left vPC (black circle) survived correction for multiple comparisons (Monte Carlo $p = 0.037$; threshold for visualization: >10 contiguous vertices at uncorrected $p < 0.005$). ***F***, Time course of the HERs (± SEM) in the left vPC. Signal that might be residually contaminated by the cardiac artifact appears in lighter color (before 80 ms). Black bar represents the time window of the significant HER difference at the sensor level. The average neural currents in this time window differed from 0 for "high" ratings ($p = 0.0017$), but not for "low" ratings ($p = 0.56$, Bayes factor = 1.78), showing that an HER could be detected in the vPC only when the self was the subject of the ongoing thought. *$p < 0.05$. **$p < 0.01$. ***$p < 0.005$.

medial frontal sensors and in a different time window (94–169 ms after T peak) (Fig. 3*A*,*B*). The cluster amplitude for trials rated as "low" or "high" corresponded to fluctuations around a baseline resting-state value (Fig. 3*C*; "high": 9.3 ± 1.9 fT, baseline: 3.4 ± 1.2 fT, "low": −5.7 ± 2.1 fT; paired *t* test, "high" vs baseline: $t_{(15)} = 2.51$, $p = 0.048$; "low" vs baseline: $t_{(15)} = -3.45$, $p = 0.0072$, Bonferroni corrected for the two tests against baseline).

As opposed to the "I" and "Me" scales, no significant difference in HERs was found between "high" and "low" Time or Valence ratings (both Monte Carlo $p > 0.33$).

Distinction between the "I" and "Me" dimensions

The differential HERs of the "Me" and "I" dimensions thus appear spatially and temporally distinct; however, the corresponding behavioral ratings were correlated. Thus, we ran two additional analyses to investigate the distinction between the "I" and "Me" dimensions.

The "I" and "Me" ratings could be capturing a general unitary self-relatedness of thoughts, as suggested by the behavioral correlation. The difference in neural correlates would in this case mostly stem from rating inaccuracy. In this view, a measure combining the ratings on the two scales should better capture the two neural correlates identified while suppressing potential noise due to rating inaccuracy. We projected the "I" and "Me" ratings on the principal component of the two scales to create a single "Self" scale, and classified trials as "high self" or "low self" relative to the median. We used the same cluster-based permutation procedure as used on the separate "I" and "Me" ratings but found no significant neural difference related to the "Self" scale (all Monte Carlo $p > 0.24$).

We then tested whether the neural correlates of the "I" and "Me" dimensions identified by the median-split approach can be attributed to the unique variance of the corresponding scale. We explored the relationship between the heartbeat-by-heartbeat cluster amplitude and the raw self-related rating at each probed thought, using a GLM with the ratings on the four scales as regressors (all variance inflation factors ≤3.52), where the regressors were orthogonalized to separate shared from unique variance. The β values corresponding to the unique variance of the "I" regressor significantly differed from 0 in the "I" cluster (GLM model 1, mean $\beta = -0.53 \pm 0.14$, *t* test against 0, one-tailed, $t_{(15)} = -3.70$, $p = 0.0011$). The β values corresponding to the unique variance of the "Me" regressor significantly differed from 0 in the "Me" cluster (GLM model 2, mean $\beta = 0.30 \pm 0.14$, *t* test against 0, one-tailed, $t_{(15)} = 2.14$, $p = 0.025$). The GLM analysis thus reveals that each self-related scale includes, in addition to shared variance revealed by the correlation between the two scales, a unique variance that covaries with neural responses to heartbeats, at distinct latencies and spatial locations.

The two control analyses thus favor the hypothesis that, even if behavioral ratings on each self-dimension are correlated, neural responses to heartbeats are preferentially associated with each self-dimension at different timings and locations.

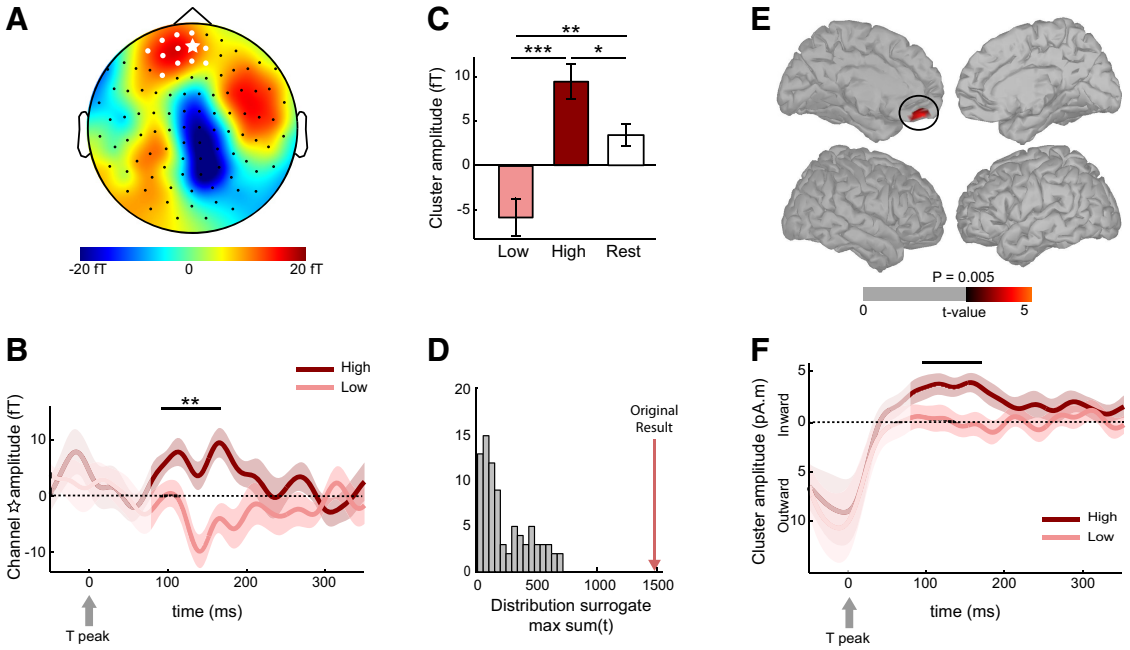


Figure 3. Differential HERs for "high" and "low" ratings on the "Me" scale. ***A***, Topographical map of the HER difference between "high" and "low" ratings on the "Me" scale, grand-averaged across 16 participants, in the 94–169 ms time window in which a significant difference was observed (Monte Carlo $p = 0.0097$, corrected for multiple comparisons). White dots represent the sensors contributing to the significant cluster. ***B***, Time course of the HER (± SEM) for "high" and "low" ratings on the "Me" scale at the sensor indicated in ***A*** (white star). The signal that might be residually contaminated by the cardiac artifact appears in lighter color (before 80 ms, not included in the analysis). Black bar represents the time window in which a significant difference was observed. ***C***, HER cluster amplitude, during thoughts rated as "high" or "low" along the "Me" scale, and during a separate eyes-open resting state session. Cluster amplitude during rest was intermediate between cluster amplitude during thoughts rated as "high" ($p = 0.048$) and cluster amplitude during thoughts rated as "low" ($p = 0.0072$). ***D***, Histogram of the distribution of the maximal cluster *t* statistic (difference between "high" and "low" trials) obtained for 100 permutations of surrogate heartbeats. The original cluster *t* statistic (arrow) lies outside the distribution of statistics obtained on surrogate data. ***E***, Neural sources of the differential HERs for thoughts rated as "high" or "low" on the "Me" scale. Only the left vmPFC (black circle) survived correction for multiple comparisons (Monte Carlo $p = 0.030$; threshold for visualization: >10 contiguous vertices at uncorrected $p < 0.005$). ***F***, Time course of the HERs (± SEM) in the left vmPFC. Signal that might be residually contaminated by the cardiac artifact appears in lighter color (before 80 ms). Black bar represents the time window of the significant HER difference at the sensor level. The average neural currents in this time window differed from 0 for "high" ratings ($p = 1.7 \times 10^{-4}$), but not for "low" ratings ($p = 1$, Bayes factor = 4.40), showing that an HER could be detected in the vmPFC only when the self was the object of the ongoing thought. *$p < 0.05$. **$p < 0.01$. ***$p < 0.005$.

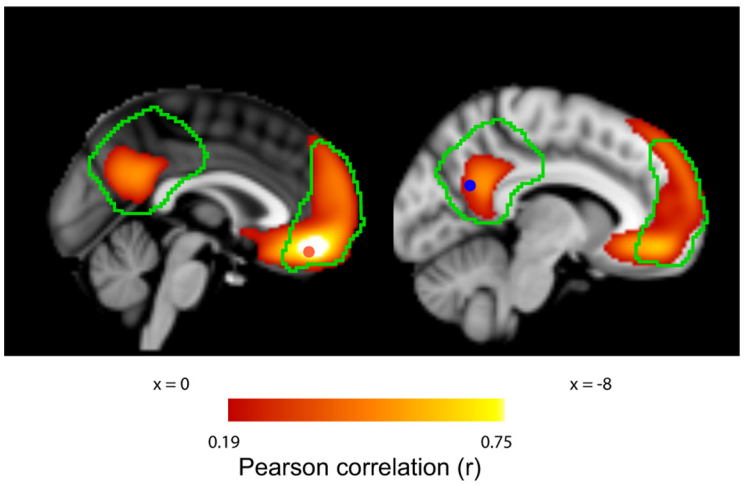


Figure 4. Functional connectivity between vmPFC and vPC and overlap with default network (DN). Red-white represents functional connectivity computed from resting-state BOLD time series of 1000 subjects at rest (Yarkoni et al., 2011), with a seed placed in left vmPFC (MNI coordinates: 0, 45, −15, left, red dot) where a differential HER along the "Me" dimension was observed. The left vPC region showing a differential HER along the "I" dimension (MNI coordinates: −8, −59, 25; right, blue dot) is functionally connected to left vmPFC (Pearson correlation $r = 0.47$). Green outline represents the DN (Laird et al., 2011).

The effects are localized in the midline regions of the DN

To identify the brain regions exhibiting distinct HERs depending on self-relatedness, we reconstructed HER sources in "high" and "low" trials on the "I" and "Me" scales, averaged the reconstructed neural currents in the time windows where significant effects were identified at the sensor level and performed a cluster-based permutation test to identify the regions that significantly contributed to the difference between "high" and "low" ratings.

HERs differed significantly along the "I" scale in the left ventral precuneus (vPC) (Fig. 2*E*,*F*; sum(t) = −93.93, Monte Carlo $p = 0.037$). The significant cluster peaked at MNI coordinates −8, −59, 25 (peak $t = -4.5$; cluster surface 4.70 cm^2). According to the AAL atlas (Tzourio-Mazoyer et al., 2002), it was centered on the left precuneus and extended dorsally and posteriorly to the cuneus and calcarine sulcus. The right homolog vPC region was also found to be responding differently but did not survive the strict correction for multiple comparisons applied here (right vPC: sum(t) = 68.20, Monte Carlo $p = 0.076$). Because the amplitude of source activity directly reflects neural currents, with the sign corresponding to current flow direction, we further tested source activity against 0, to find out in which condition heartbeats elicited a detectable neural response. In the left vPC, source activity

Table 2. Cardiorespiratory parameters and arousal-related measures (mean ± SEM) do not differ between "high" and "low" ratings on either the "I" or the "Me" scale[a]

	"I" scale				"Me" scale			
	Mean "high"	Mean "low"	Paired *t* test (uncorrected *p*)	Bayes factor	Mean "high"	Mean "low"	Paired *t* test (uncorrected *p*)	Bayes factor
Interbeat interval (ms)	941 ± 21	945 ± 21	$t_{(15)} = -1.56$, $p = 0.14$	Inconclusive (1.13)	943 ± 22	943 ± 21	$t_{(15)} = 0.06$, $p = 0.95$	Substantial (4.39)
Heart rate variability (ms)	59 ± 6.7	59 ± 7.0	$t_{(15)} = -0.048$, $p = 0.96$	Substantial (4.40)	59 ± 6.2	59 ± 7.4	$t_{(15)} = -0.13$, $p = 0.90$	Substantial (4.35)
Respiratory cycle duration (s)	3.65 ± 0.13	3.69 ± 0.13	$t_{(15)} = -0.51$, $p = 0.62$	Substantial (3.58)	3.67 ± 0.12	3.67 ± 0.12	$t_{(15)} = 0.21$, $p = 0.83$	Substantial (4.25)
Respiratory phase difference (phase bifurcation index (PBI) against 0)	Mean PBI = $-1.60 \times 10^{-4} \pm 3.90 \times 10^{-4}$		$t_{(15)} = -0.40$, $p = 0.69$	Substantial (3.87)	Mean PBI = $2.20 \times 10^{-5} \pm 1.60 \times 10^{-4}$		$t_{(15)} = 0.14$, $p = 0.89$	Substantial (4.34)
Pupil diameter (a.u.)	−17 ± 31	−19 ± 40	$t_{(15)}$=0.028, $p = 0.98$	Substantial (4.40)	−6.7 ± 22	−30 ± 35	$t_{(15)} = 0.44$, $p = 0.67$	Substantial (3.78)
Electrodermal activity (a.u.)	$-8.20 \times 10^{-3} \pm 0.042$	0.026 ± 0.034	$t_{(15)} = -0.46$, $p = 0.65$	Substantial (3.72)	0.025 ± 0.036	$-6.00 \times 10^{-3} \pm 0.032$	$t_{(15)} = 0.48$, $p = 0.64$	Substantial (3.68)
Mean alpha power (occipitoparietal sensors, $fT^2\,Hz^{-1}$)	$8.10 \times 10^5 \pm 1.60 \times 10^5$	$8.60 \times 10^5 \pm 2.00 \times 10^5$	$t_{(15)} = -0.84$, $p = 0.41$	Inconclusive (2.57)	$8.20 \times 10^5 \pm 1.80 \times 10^5$	$8.50 \times 10^5 \pm 1.80 \times 10^5$	$t_{(15)} = -0.59$, $p = 0.57$	Substantial (3.36)
No. of blinks	3.1 ± 1.26	3.2 ± 1.19	$t_{(15)} = -0.25$, $p = 0.81$	Substantial (4.19)	3.0 ± 1.35	3.3 ± 1.16	$t_{(15)} = -0.32$, $p = 0.75$	Substantial (4.05)
No. of small saccades (<2 degrees)	89.6 ± 15.49	85.0 ± 15.20	$t_{(15)} = 0.90$, $p = 0.38$	Inconclusive (2.40)	84.3 ± 14.41	90.3 ± 16.64	$t_{(15)} = -0.82$, $p = 0.43$	Inconclusive (2.65)

[a]Bayes factors quantify the amount of evidence in favor of the absence of a difference between "high" and "low" ratings. PBI

significantly differed from 0 for "high" ratings on the "I" scale ("high": −2.03 ± 0.50 pA.m, *t* test against 0: $t_{(15)} = -4.16$, $p = 0.0017$, Bonferroni corrected for the two comparisons against baseline) but did not differ from 0 in thoughts rated as "low" on the "I" scale ("low": 0.57 ± 0.52 pA.m, $t_{(15)} = 1.12$, $p = 0.56$, Bonferroni corrected, Bayes factor = 1.78). A HER can be detected in the left vPC only when the self is the subject of the ongoing thought.

The differential HERs along the "Me" scale were located in the left ventromedial prefrontal cortex (vmPFC) (Fig. 3*E*,*F*; sum(*t*) = −93.94, Monte Carlo $p = 0.030$). The significant cluster peaked at MNI coordinates 0, 45, −15 (peak $t = -4.7$; cluster surface 4.37 cm^2), was centered on the left frontal medial orbital gyrus and extended posteriorly and dorsally to the left anterior cingulate and rectus gyri. A HER could be detected in the vmPFC only when the self is the object of the ongoing thought (mean neural current in the left vmPFC, *t* test against 0, Bonferroni corrected for the two comparisons against baseline, "high": −3.69 ± 0.72 pA.m, $t_{(15)} = -5.32$, $p = 1.7 \times 10^{-4}$, "low": −0.040 ± 0.86 pA.m, $t_{(15)} = -0.05$, $p = 1$, Bayes factor = 4.40).

To compare these locations with results from fMRI resting-state connectivity (Fig. 4), we superimposed our results with the DN, as described Laird et al. (2011). The two regions differentially activated by heartbeats are indeed part of the DN. We further verified, based on resting connectivity maps in 1000 subjects (Yarkoni et al., 2011), that the two regions differentially responding to heartbeats are functionally connected at rest (Pearson correlation $r = 0.47$ between resting-state fMRI time series at MNI coordinates 0, 45, −15 and −8, −59, 25).

Cardiorespiratory and arousal measures do not vary with self-relatedness

The effects reported here are not trivially explained by massive changes in bodily state along the "I" or "Me" scales. There was no sign that cardiac activity differed between "high" and "low" ratings, on the cardiorespiratory parameters we measured (interbeat interval, heart rate variability, respiratory cycle duration, respiratory phase, all uncorrected $p \geq 0.14$, all but one of the Bayes factors were ≥3.58, indicating substantial evidence for the absence of an effect; Table 2). We further verified that there was no difference between "high" and "low" ratings for a number of arousal-related measures (Luft and Bhattacharya, 2015) (Table 2) on both self-related scales: electrodermal activity (both $p \geq 0.64$, Bayes factors ≥3.68), pupil diameter (Fig. 5*B*; both $p \geq 0.67$, Bayes factors ≥3.78), alpha power (Fig. 5*C*; 8–12 Hz, averaged over occipitoparietal sensors, both $p \geq 0.41$, Bayes factors ≥2.57). Last, the number of blinks (both $p \geq 0.75$, Bayes factor ≥4.05) and saccades (both $p \geq 0.38$, Bayes factors ≥2.40) did not vary either (Table 2).

Control: the effects are of neural origin and time-locked to heartbeats

To show that the observed effects were truly locked to heartbeats and not driven by slow fluctuations of neural activity distinguishing between "high" and "low" ratings, we created for each participant 100 permutations of surrogate heartbeats with the same interbeat intervals as in the original data. We then compared surrogate HERs for "high" and "low" trials and computed the largest cluster *t* statistic at each permutation. None of the 100 permutations generated a cluster *t* statistic as large as the ones originally obtained with heartbeat-locked data (Figs. 2*D*, 3D); thus, the differential effects reported here appear to be truly locked to heartbeats (Monte Carlo $p < 0.01$).

We then analyzed the ECG, to check that the effects observed on MEG data were not reflecting a difference in the electrical activity of the heart directly picked up by the MEG sensors. The ECG was recorded from seven electrodes around the base of the neck (vertical leads) and seven horizontal derivations between neighboring electrodes were computed offline. The ECG appeared similar in "high" versus "low" trials on both scales (Fig. 5*A*). The same cluster-based permutation test as used on MEG sensors applied to ECG data did not generate any candidate cluster, neither for the "I" nor for the "Me" scale, and neither on vertical nor horizontal ECG leads. Testing the time windows for which we obtained significant differences in MEG activity revealed no difference in the ECG signal, on either scale (paired *t* test, "high" vs "low"; "I" scale, mean ECG amplitude averaged between 298 and 327 ms after T peak at each vertical or horizontal derivation, all $|t_{(15)}| < 0.77$, all uncorrected $p > 0.46$, all Bayes

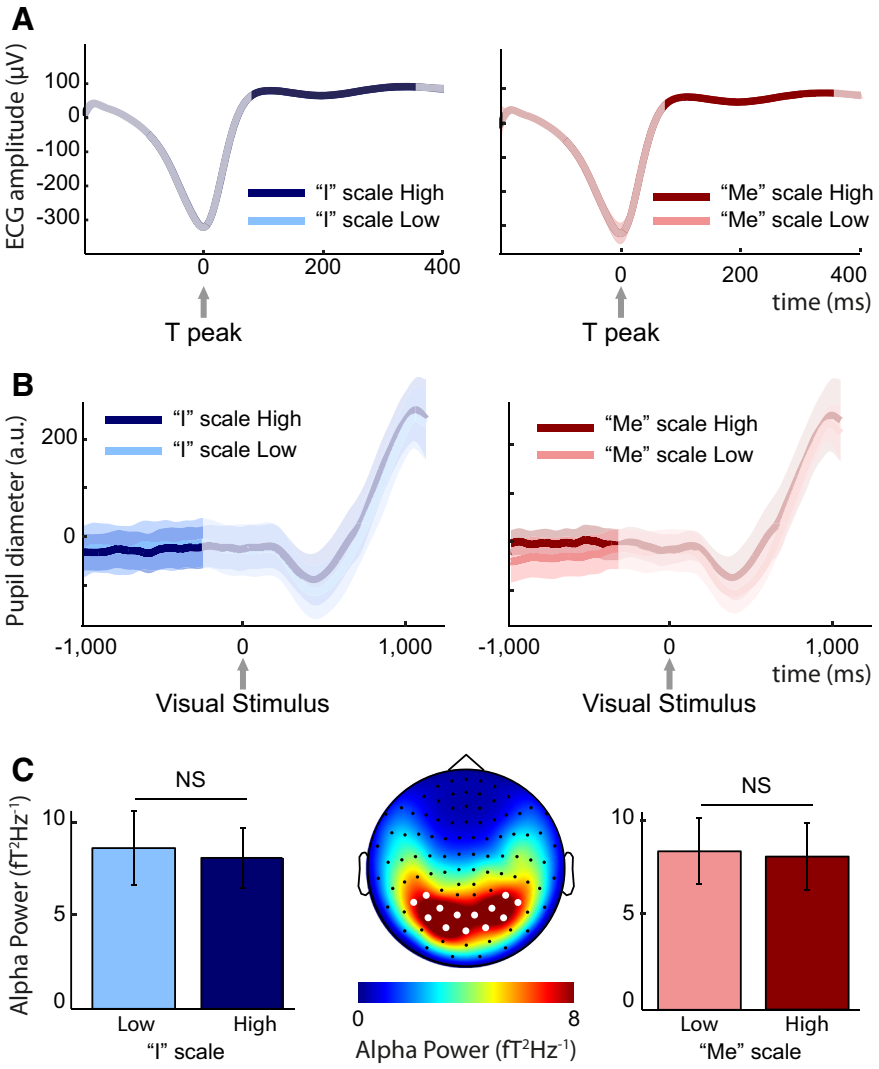


Figure 5. Controls. ***A***, Time course of the average ECG signal (± SEM) for "high" and "low" ratings along the "I" (left) and "Me" (right) scales, on the vertical derivation lead II. The signal appearing in darker color corresponds to the time window that was analyzed in the MEG data. The ECG, recorded from seven electrodes around the base of the neck to carefully monitor the potential direct contribution of heart electrical activity to MEG signals, appeared similar in "high" versus "low" trials on both scales, and no significant differences were found. ***B***, Time course of the average pupil diameter (± SEM) signal for "high" and "low" ratings along the "I" (left) and "Me" (right) scales. We analyzed the time window during the thought, from the last-but-one heartbeat to 400 ms preceding the visual stimulus (signal in darker color). We observed no statistical difference between "high" and "low" trials for either the "I" or the "Me" scales (both $p \geq 0.67$, both Bayes factors ≥ 3.78). ***C***, Average alpha power (± SEM) for "high" and "low" ratings along the "I" (left) and "Me" (right) scales, on the 15 sensors with the largest alpha power across conditions, indicated by white dots in the alpha power topographical map (center). We did not observe differences in alpha power between "high" and "low" trials for either scale (both $p \geq 0.41$, both Bayes factors ≥ 2.57). NS, Not significant.

factors ≥ 2.81; "Me" scale, mean amplitude averaged between 94 and 169 ms at each derivation, all $|t_{(15)}| < 0.95$, all uncorrected $p > 0.36$, all Bayes factors ≥ 2.25).

Intersubject variability in various personality traits or interoceptive abilities did not contribute to the effects

We tested whether the individual amplitude of the effects (e.g., the cluster amplitude difference between "high" and "low" ratings) correlated with a number of personality aspects (self-consciousness scale, daydreaming frequency scale, trait anxiety inventory) or interoceptive ability as measured in the heartbeat counting task. None of these measures correlated with the amplitude of the effects on either scale over participants (Table 3).

Discussion

Our results reveal a direct link between selfhood and neural responses to heartbeats in the DN. We show that self-relatedness is parametrically encoded in neural responses to heartbeats in two midline regions of the DN that have been repeatedly associated with the self in the fMRI literature (Qin and Northoff, 2011). We verified that the neural events we describe are locked to heartbeats and cannot be due to the cardiac field artifact. More generally, we could not measure any significant changes in cardiorespiratory parameters (heart rate, heart rate variability, respiration rate, or phase) or in classical measures of arousal (electrodermal activity, pupil diameter, alpha rhythm power). Our findings indicate that the two seemingly distinct roles of the DN, in self-related cognition (Buckner et al., 2008; Qin and Northoff, 2011; Andrews-Hanna et al., 2014) on the one hand, and in the monitoring of bodily signal for autonomous function regulation (Thayer et al., 2012; Beissner et al., 2013) on the other, are functionally coupled. The two novel self-related scales that we developed enable us to further specify the functional role of the vPC and vmPFC, that appear to relate to the "I" and the "Me" aspects of the self, respectively.

Specifying the respective roles of the vPC and vmPFC: distinguishing between the "I" and the "Me"

We find that HERs in vPC and vmPFC covary preferentially with the "I" and "Me" dimensions of the self, respectively. The "I" and "Me" distinction is only partial because the corresponding ratings were behaviorally correlated, but we verified that a general self-relatedness measure combining the two scales together did not reproduce the results and, conversely, that the results presented here can be accounted for by the variance unique to each self-related scale. Our results thus suggest that the conceptual distinction originally proposed by James (1890) between the "I" and the "Me" has some biological counterpart and provides a useful theoretical framework to specify the respective roles of vPC and vmPFC.

The "I" is prereflective in the sense that it refers to the subject who is experiencing something from the first-person perspective, without necessarily reflecting on the experience itself (Legrand and Ruby, 2009; Christoff et al., 2011; Gallagher, 2012). The implicit and pervasive "I" is possibly the most basic aspect of the self, yet little is known about its specific neural correlates (Christoff et al., 2011). Our results indicate that vPC is preferentially related to the "I." A closer look at the literature indicates that the vPC is active in tasks as diverse as episodic memory retrieval (Martinelli

Table 3. Scores on personality trait questionnaires and interoceptive abilities do not correlate with the amplitude of the cluster difference between "high" and "low" ratings[a]

		"I" cluster		"Me" cluster	
	Scores (mean ± SEM)	Pearson correlation (*p* values, Bonferroni corrected for the 4 scales)	Bayes factor on regression	Pearson correlation (*p* values, Bonferroni corrected for the 4 scales)	Bayes factor on regression
Self-consciousness scale	38.06 ± 2.68	$r_{(14)} = -0.37, p = 0.64$	Inconclusive (1.15)	$r_{(14)} = -0.37, p = 0.63$	Inconclusive (1.14)
Daydreaming frequency scale	43.25 ± 1.65	$r_{(14)} = -0.36, p = 0.68$	Inconclusive (1.17)	$r_{(14)} = -0.33, p = 0.84$	Inconclusive (1.33)
State-trait anxiety inventory (trait inventory)	37.19 ± 1.90	$r_{(14)} = -0.43, p = 0.40$	Inconclusive (1.14)	$r_{(14)} = -0.23, p = 1$	Inconclusive (1.80)
Interoceptive abilities	0.79 ± 0.025	$r_{(14)} = -0.053, p = 1$	Inconclusive (2.31)	$r_{(14)} = -0.089, p = 1$	Inconclusive (2.25)

[a]Bayes factors were computed on the regression to quantify the amount of evidence in favor of the absence of an effect.

et al., 2013), perspective taking (Vogeley et al., 2004; Cavanna and Trimble, 2006), body ownership, self-location (Guterstam et al., 2015), spatial navigation, imagination, and future planning (Vann et al., 2009), and the feeling of agency (Miele et al., 2011; Nahab et al., 2011), all of which require the underlying and implicit engagement of the self as the subject. Conversely, the "Me" involves the explicit reflection about oneself and appears here more particularly linked to vmPFC. Self-attribution of personality traits (i.e., a task that particularly involves the "Me") recruits preferentially medial prefrontal structures (Martinelli et al., 2013). Our results thus suggest a refined interpretation of the self-processing literature in terms of the self being the subject of experience or the object of reflection. This relates to a more general debate on the distinction between experiencing and introspecting about experience that is beginning to receive some attention-notably in the literature on conscious vision (Frässle et al., 2014) and agency (Synofzik et al., 2008).

Functional coupling between physiological monitoring and self-related processing in the DN

Our results show a systematic covariation, down to the level of single trials, between ratings of self-relatedness and the amplitude of neural responses to heartbeats in the DN. The whole-brain approach used here did not reveal differential neural responses to heartbeats outside the DN, notably in the insula. Our results indicate that the two roles of the DN, namely, physiological monitoring and self-related processing, are not merely colocalized but are functionally coupled and thus should be considered in the same functional framework.

The vmPFC is a known visceral monitoring center (Vogt and Derbyshire, 2009) previously found to respond to heartbeats in the same latency range (Park et al., 2014). Although the vPC is not a direct target of visceral inputs, it is functionally connected to visceral centers of the brain (Zhang and Li, 2012) and it is involved in autonomous functions (Beissner et al., 2013). vPC may therefore be receiving visceral information through one or more cortical relays, which is compatible with the longer latency of the effect observed in vPC. It is difficult to infer from our data whether and how the latency difference in transient neural responses to heartbeats in vPC and vmPFC directly relate to a differential time course of the "I" and "Me" dimensions in spontaneous thought that probably develop over seconds. This issue directly relates to the general and challenging question of the temporal mapping between neural events and mental events. For instance, in vision, it is known that different attributes of the same object, such as color or motion, are neurally processed at different speeds. Whether and how different neural processing speeds are compensated for, or contribute to the final percept, is still a debated issue.

The functional coupling between HERs and self-relatedness could stem from different mechanisms. As presented in the Introduction, theories grounding the self into an integrated neural map of the organism (Damasio, 1999; Craig, 2009; Park and Tallon-Baudry, 2014) would predict that HERs directly contribute to the specification of the self. HERs would contribute to the constant update of a neural reference frame centered on the subject's body that would serve as a basis for the development of self-relatedness. Our results directly support these theories; however, other interpretations should be considered. Self-related thoughts could induce an internally directed attentional shift, thereby amplifying the processing of internal signals, including heartbeats (Montoya et al., 1993). Explicitly orienting attention toward heartbeats alters activity in the insula, somatomotor, and dorsal anterior cingulate cortices (Critchley et al., 2004; Pollatos et al., 2005, 2007; Canales-Johnson et al., 2015). None of these regions showed differential activation in the present experiment, making an attentional account of our results unlikely. One could also argue that HER covariation with self-relatedness is a byproduct of self-related processing, with neurons responding to heartbeats being modulated by neurons encoding self-relatedness. In this view, HERs are modulated by the self-relatedness of spontaneous thoughts but have no direct consequence on the contents of those thoughts. Determining whether HER modulations are a mere byproduct of self-relatedness or play an active role in the construction of selfhood amounts to moving from correlation to causation, a notoriously difficult achievement.

Our results are coherent with the large body of fMRI evidence revealing the role of the DN in self-related processing and spontaneous cognition but call for a reappraisal of the importance of physiological monitoring in the DN (Iacovella and Hasson, 2011). While covariations of brain activity and peripheral measures of autonomic functions have often been dismissed as mere "physiological noise," which should be regressed out of the data (Glover et al., 2000; Shmueli et al., 2007; Birn, 2012), there is now converging evidence that the DN is truly engaged in physiological regulation (Nagai et al., 2004; Fan et al., 2012; Thayer et al., 2012; Beissner et al., 2013; Chang et al., 2013). This physiological view accounts well for a number of facts about the DN that are not always easily explained by self-oriented cognition, such as the high basal metabolic rate of the DN (Minoshima et al., 1997), its persistence in early sleep stages (Horovitz et al., 2008; Larson-Prior et al., 2009) and light sedation (Greicius et al., 2008), or its conservation across species (Mantini et al., 2011; Lu et al., 2012). It has been argued that the implication of the DN in general physiological or "maintenance" functions speaks against a specific cognitive role of the DN (Larson-Prior et al., 2009). On the contrary, our results show that, even in the absence of bodily changes as indexed by classical peripheral measures, neural responses to heartbeats in the DN encode cognitively refined infor-

mation about the self. This implies that physiological and cognitive functions should be considered jointly in the DN.

References

Andrews-Hanna JR, Smallwood J, Spreng RN (2014) The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann N Y Acad Sci 1316:29–52. CrossRef Medline

Armour JA, Ardell JL (2004) Basic and clinical neurocardiology (Armour JA, Ardell JL, eds). New York: Oxford UP.

Beissner F, Meissner K, Bär KJ, Napadow V (2013) The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. J Neurosci 33:10503–10511. CrossRef Medline

Birn RM (2012) The role of physiological noise in resting-state functional connectivity. Neuroimage 62:864–870. CrossRef Medline

Bishara AJ, Hittner JB (2012) Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches. Psychol Methods 17:399–417. CrossRef Medline

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38. CrossRef Medline

Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. J Neurosci 29:7869–7876. CrossRef Medline

Canales-Johnson A, Silva C, Huepe D, Rivera-Rei Á, Noreika V, Garcia Mdel C, Silva W, Ciraolo C, Vaucheret E, Sedeño L, Couto B, Kargieman L, Baglivo F, Sigman M, Chennu S, Ibáñez A, Rodríguez E, Bekinschtein TA (2015) Auditory feedback differentially modulates behavioral and neural markers of objective and subjective performance when tapping to your heartbeat. Cereb Cortex 25:4490–4503. CrossRef Medline

Cavanna AE, Trimble MR (2006) The precuneus: a review of its functional anatomy and behavioural correlates. Brain 129:564–583. CrossRef Medline

Chang C, Metzger CD, Glover GH, Duyn JH, Heinze HJ, Walter M (2013) Association between heart rate variability and fluctuations in resting-state functional connectivity. Neuroimage 68:93–104. CrossRef Medline

Christoff K, Cosmelli D, Legrand D, Thompson E (2011) Specifying the self for cognitive neuroscience. Trends Cogn Sci 15:104–112. CrossRef Medline

Couto B, Adolfi F, Velasquez M, Mesow M, Feinstein J, Canales-Johnson A, Mikulan E, Martínez-Pernía D, Bekinschtein T, Sigman M, Manes F, Ibanez A (2015) Heart evoked potential triggers brain responses to natural affective scenes: a preliminary study. Auton Neurosci 193:132–137. CrossRef Medline

Craig AD (2009) How do you feel now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70. CrossRef Medline

Critchley HD, Harrison NA (2013) Visceral influences on brain and behavior. Neuron 77:624–638. CrossRef Medline

Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ (2004) Neural systems supporting interoceptive awareness. Nat Neurosci 7:189–195. CrossRef Medline

Damasio AR (1999) The feeling of what happens: body, emotion and the making of consciousness. San Diego: Harcourt.

D'Argembeau A, Feyers D, Majerus S, Collette F, Van der Linden M, Maquet P, Salmon E (2008) Self-reflection across time: cortical midline structures differentiate between present and past selves. Soc Cogn Affect Neurosci 3:244–252. CrossRef Medline

Dirlich G, Dietl T, Vogl L, Strian F (1998) Topography and morphology of heart action-related EEG potentials. Electroencephalogr Clin Neurophysiol 108:299–305. CrossRef Medline

Fan J, Xu P, Van Dam NT, Eilam-Stock T, Gu X, Luo YJ, Hof PR (2012) Spontaneous brain activity relates to autonomic arousal. J Neurosci 32:11176–11186. CrossRef Medline

Fenigstein A, Scheier MF, Buss AH (1975) Public and private self-consciousness: assessment and theory. J Consult Clin Psychol 43:522–527. CrossRef

Fischl B, van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically parcellating the human cerebral cortex. Cereb Cortex 14:11–22. CrossRef Medline

Frässle S, Sommer J, Jansen A, Naber M, Einhäuser W (2014) Binocular rivalry: frontal activity relates to introspection and action but not to perception. J Neurosci 34:1738–1747. CrossRef Medline

Gallagher S (2012) Phenomenology. New York: Macmillan.

Giambra LM (1993) The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. Exp Gerontol 28:485–492. CrossRef Medline

Glover GH, Li TQ, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn Reson Med 44:162–167. CrossRef Medline

Gray MA, Taggart P, Sutton PM, Groves D, Holdright DR, Bradbury D, Brull D, Critchley HD (2007) A cortical potential reflecting cardiac function. Proc Natl Acad Sci U S A 104:6818–6823. CrossRef Medline

Greicius MD, Kiviniemi V, Tervonen O, Vainionpää V, Alahuhta S, Reiss AL, Menon V (2008) Persistent default-mode network connectivity during light sedation. Hum Brain Mapp 29:839–847. CrossRef Medline

Guterstam A, Björnsdotter M, Gentile G, Ehrsson HH (2015) Posterior cingulate cortex integrates the senses of self-location and body ownership. Curr Biol 25:1416–1425. CrossRef Medline

Horovitz SG, Fukunaga M, de Zwart JA, van Gelderen P, Fulton SC, Balkin TJ, Duyn JH (2008) Low frequency BOLD fluctuations during resting wakefulness and light sleep: a simultaneous EEG-fMRI study. Hum Brain Mapp 29:671–682. CrossRef Medline

Hurlburt RT, Heavey CL (2001) Telling what we know: describing inner experience. Trends Cogn Sci 5:400–403. CrossRef Medline

Iacovella V, Hasson U (2011) The relationship between BOLD signal and autonomic nervous system functions: implications for processing of "physiological noise." Magn Reson Imaging 29:1338–1345. CrossRef Medline

James W (1890) The principles of psychology. New York: Holt.

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795. CrossRef Medline

Kern M, Aertsen A, Schulze-Bonhage A, Ball T (2013) Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. Neuroimage 81:178–190. CrossRef Medline

Laird AR, Fox PM, Eickhoff SB, Turner JA, Ray KL, McKay DR, Glahn DC, Beckmann CF, Smith SM, Fox PT (2011) Behavioral interpretations of intrinsic connectivity networks. J Cogn Neurosci 23:4022–4037. CrossRef Medline

Larson-Prior LJ, Zempel JM, Nolan TS, Prior FW, Snyder AZ, Raichle ME (2009) Cortical network functional connectivity in the descent to sleep. Proc Natl Acad Sci U S A 106:4489–4494. CrossRef Medline

Lechinger J, Heib DP, Gruber W, Schabus M, Klimesch W (2015) Heartbeat-related EEG amplitude and phase modulations from wakefulness to deep sleep: interactions with sleep spindles and slow oscillations. Psychophysiology 52:1441–1450. CrossRef Medline

Legrand D, Ruby P (2009) What is self-specific? Theoretical investigation and critical review of neuroimaging results. Psychol Rev 116:252–282. CrossRef Medline

Liang F, Paulo R, Molina G, Clyde MA, Berger JO (2008) Mixtures of g priors for Bayesian variable selection. J Am Stat Assoc 103:410–423. CrossRef

Lu H, Zou Q, Gu H, Raichle ME, Stein EA, Yang Y (2012) Rat brains also have a default mode network. Proc Natl Acad Sci U S A 109:3979–3984. CrossRef Medline

Luft CD, Bhattacharya J (2015) Aroused with heart: modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. Sci Rep 5:15717. CrossRef Medline

Mantini D, Gerits A, Nelissen K, Durand JB, Joly O, Simone L, Sawamura H, Wardak C, Orban GA, Buckner RL, Vanduffel W (2011) Default mode of brain function in monkeys. J Neurosci 31:12954–12962. CrossRef Medline

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190. CrossRef Medline

Martinelli P, Sperduti M, Piolino P (2013) Neural substrates of the self-memory system: new insights from a meta-analysis. Hum Brain Mapp 34:1515–1529. CrossRef Medline

Miele DB, Wager TD, Mitchell JP, Metcalfe J (2011) Dissociating neural correlates of action monitoring and metacognition of agency. J Cogn Neurosci 23:3620–3636. CrossRef Medline

Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE (1997) Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Ann Neurol 42:85–94. CrossRef Medline

Montoya P, Schandry R, Müller A (1993) Heartbeat evoked potentials

(HEP): topography and influence of cardiac awareness and focus of attention. Electroencephalogr Clin Neurophysiol 88:163–172. CrossRef Medline

Nagai Y, Critchley HD, Featherstone E, Trimble MR, Dolan RJ (2004) Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: a physiological account of a "default mode" of brain function. Neuroimage 22:243–251. CrossRef Medline

Nahab FB, Kundu P, Gallea C, Kakareka J, Pursley R, Pohida T, Miletta N, Friedman J, Hallett M (2011) The neural processes underlying self-agency. Cereb Cortex 21:48–55. CrossRef Medline

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869. CrossRef Medline

Park HD, Tallon-Baudry C (2014) The neural subjective frame: from bodily signals to perceptual consciousness. Philos Trans R Soc Lond B Biol Sci 369:20130208. CrossRef Medline

Park HD, Correia S, Ducorps A, Tallon-Baudry C (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat Neurosci 17:612–618. CrossRef Medline

Pollatos O, Kirsch W, Schandry R (2005) Brain structures involved in interoceptive awareness and cardioafferent signal processing: a dipole source localization study. Hum Brain Mapp 26:54–64. CrossRef Medline

Pollatos O, Schandry R, Auer DP, Kaufmann C (2007) Brain structures mediating cardiovascular arousal and interoceptive awareness. Brain Res 1141:178–187. CrossRef Medline

Powell LJ, Macrae CN, Cloutier J, Metcalfe J, Mitchell JP (2010) Dissociable neural substrates for agentic versus conceptual representations of self. J Cogn Neurosci 22:2186–2197. CrossRef Medline

Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? Neuroimage 57:1221–1233. CrossRef Medline

Schandry R (1981) Heart beat perception and emotional experience. Psychophysiology 18:483–488. CrossRef Medline

Schandry R, Montoya P (1996) Event-related brain potentials and the processing of cardiac activity. Biol Psychol 42:75–85. CrossRef Medline

Schandry R, Sparrer B, Weitkunat R (1986) From the heart to the brain: a study of heartbeat contingent scalp potentials. Int J Neurosci 30:261–275. CrossRef Medline

Shmueli K, van Gelderen P, de Zwart JA, Horovitz SG, Fukunaga M, Jansma JM, Duyn JH (2007) Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. Neuroimage 38:306–320. CrossRef Medline

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists.

Synofzik M, Vosgerau G, Newen A (2008) Beyond the comparator model: a multifactorial two-step account of agency. Conscious Cogn 17:219–239. CrossRef Medline

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011:879716. CrossRef Medline

Thayer JF, Ahs F, Fredrikson M, Sollers JJ 3rd, Wager TD (2012) A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neurosci Biobehav Rev 36:747–756. CrossRef Medline

Tusche A, Smallwood J, Bernhardt BC, Singer T (2014) Classifying the wandering mind: revealing the affective content of thoughts during task-free rest periods. Neuroimage 97:107–116. CrossRef Medline

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289. CrossRef Medline

Van Eden CG, Buijs RM (2000) Functional neuroanatomy of the prefrontal cortex: autonomic interactions. Prog Brain Res 126:49–62. CrossRef Medline

Vann SD, Aggleton JP, Maguire EA (2009) What does the retrosplenial cortex do? Nat Rev Neurosci 10:792–802. CrossRef Medline

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CM (2010) The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. Neuroimage 51:112–122. CrossRef Medline

Vogeley K, May M, Ritzl A, Falkai P, Zilles K, Fink GR (2004) Neural correlates of first-person perspective as one constituent of human self-consciousness. J Cogn Neurosci 16:817–827. CrossRef Medline

Vogt BA, Derbyshire SWG (2009) Visceral circuits and cingulate-mediated autonomic functions. In: Cingulate neurobiology and disease (Vogt BA, ed), pp 219–236. Oxford: Oxford UP.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670. CrossRef Medline

Zhang S, Li CS (2012) Functional connectivity mapping of the human precuneus by resting state fMRI. Neuroimage 59:3548–3562. CrossRef Medline

VI. Article II: Is the cardiac monitoring function related to the self in both the default-network and right anterior insula?

A. Abstract in French

Les théories du soi proposent que celui-ci est ancré dans le suivi des signaux corporels par le cerveau. Le soi devrait donc être associé aux régions cérébrales intéroceptives, notamment l'insula antérieure droite. Cependant, les études sur le soi montrent le rôle des régions médiales du réseau du mode par défaut, sans faire référence au suivi des signaux viscéraux. Ici, nous avons étudié cette apparente contradiction. Nous avons montré précédemment en magnétoencéphalographie (MEG) que les réponses aux battements cardiaques dans le réseau du mode par défaut encodent deux dimensions du soi, le « Je » agentif et le « Moi » introspectif. Ici, nous confirmons et détaillons anatomiquement ce résultat avec des enregistrements intracérébraux. Nous montrons chez deux patients une corrélation entre des réponses évoquées aux battements cardiaques et le rapport au soi des pensées, essai par essai. Une analyse par région d'intérêt de l'insula montre de plus que les réponses aux battements cardiaques dans cette région, enregistrées en MEG, encodent la dimension « Je » des pensées spontanées. L'effet dans l'insula antérieure droite est plus faible que l'effet dans le réseau du mode par défaut, et n'a été répliqué en iEEG que chez un patient sur deux. Nous proposons qu'un mécanisme commun, le suivi des signaux cardiaques par le cerveau, sous-tend le soi dans le réseau du mode par défaut et l'insula antérieure droite. Ceci pourrait réconcilier les études sur le soi, en incluant le réseau du mode par défaut, et les études sur l'intéroception, qui se focalisent sur l'insula.

B. Article

PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Research

Cite this article: Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C. 2016 Is the cardiac monitoring function related to the self in both the default network and right anterior insula? *Phil. Trans. R. Soc. B* **371**: 20160004.
http://dx.doi.org/10.1098/rstb.2016.0004

Accepted: 5 August 2016

One contribution of 16 to a theme issue 'Interoception beyond homeostasis: affect, cognition and mental health'.

Subject Areas:
cognition, neuroscience

Keywords:
intracranial electroencephalography, magnetoencephalography, neural responses to heartbeats, heartbeat-evoked responses, interoception, spontaneous cognition

Author for correspondence:
Mariana Babo-Rebelo
e-mail: mariana.babo-rebelo@ens.fr

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3464844.

THE ROYAL SOCIETY PUBLISHING

Is the cardiac monitoring function related to the self in both the default network and right anterior insula?

Mariana Babo-Rebelo[1], Nicolai Wolpert[1], Claude Adam[2], Dominique Hasboun[2] and Catherine Tallon-Baudry[1]

[1]Laboratoire de Neurosciences Cognitives (ENS – INSERM U960), Département d'Etudes Cognitives, Ecole Normale Supérieure – PSL Research University, 75005 Paris, France
[2]AP-HP, Groupe Hospitalier Pitié-Salpêtrière, Paris 75013, France

MB, 0000-0001-8664-7427; CT, 0000-0001-8480-5831

The self has been proposed to be rooted in the neural monitoring of internal bodily signals and might thus involve interoceptive areas, notably the right anterior insula (rAI). However, studies on the self consistently showed the involvement of midline default network (DN) nodes, without referring to visceral monitoring. Here, we investigate this apparent discrepancy. We previously showed that neural responses to heartbeats in the DN encode two different self-dimensions, the agentive 'I' and the introspective 'Me', in a whole-brain analysis of magnetoencephalography (MEG) data. Here, we confirm and anatomically refine this result with intracranial recordings (intracranial electroencephalography, iEEG). In two patients, we show a parametric modulation of neural responses to heartbeats by the self-relatedness of thoughts, at the single trial level. A region-of-interest analysis of the insula reveals that MEG responses to heartbeats in the rAI encode the 'I' self-dimension. The effect in rAI was weaker than in the DN and was replicated in iEEG data in one patient out of two. We propose that a common mechanism, the neural monitoring of cardiac signals, underlies the self in both the DN and rAI. This might reconcile studies on the self highlighting the DN, with studies on interoception focusing on the insula.

This article is part of the themed issue 'Interoception beyond homeostasis: affect, cognition and mental health'.

1. Introduction

It has been proposed that the self is rooted in the neural monitoring of internal bodily signals [1,2]. For Damasio [1], for instance, the non-conscious cartography of bodily states, the 'proto-self', is the basis for the construction of higher level conscious forms of self, the 'core self' and the 'autobiographical self'. Experimental studies of the neural bases of visceral information processing in humans have mostly relied on explicit interoception paradigms, where attention is voluntarily oriented towards internal signals and thus towards oneself. The role of the right anterior insula (rAI) in cardiac interoception has been particularly underlined, following Craig's influential theory [3] that awareness arises from the integration of visceral signals with environmental, hedonic, motivational, social and cognitive signals, in a gradient along the insular cortex, but also based on empirical findings. Indeed, both the level of activation and grey matter volume of the rAI correlate with performance in the heartbeat-counting task [4]. An involvement of insular regions during the heartbeat-counting task [5] is also compatible with the localization of the attentional modulation of heartbeat-evoked responses (HERs) [6,7]. However, the role of the rAI in explicit interoception remains debated, because interoceptive accuracy was preserved in a patient with bilateral insula damage [8]. In addition, in the heartbeat

detection task, cardiac interoception modulates activity in a variety of other areas, such as somatomotor areas and the dorsal anterior cingulate cortex [4,9]. Most notably, the rAI is one of the structures most commonly activated across all cognitive tasks [10,11], and might play a more general role in switching between internally and externally oriented cognition [12].

Besides, most experimental studies of the self do not point at the insula, but at the default network (DN) [13], a network of brain regions that is more active at rest [14], during spontaneous thoughts [15] and internally directed cognition [16], than during most cognitively demanding tasks [17]. As shown in a meta-analysis [18], tasks pertaining to the *cognitive self*, such as autobiographical memory, self versus other personality trait judgement, own name detection or face recognition, consistently involve the medial nodes of the DN. This vast experimental literature does not make any explicit reference to the body or to the processing of bodily signals, and thus appears disconnected from theories relating the self to bodily signals. This overview of studies on the self and explicit interoception thus suggests the involvement of two sets of regions, the DN that is involved in the self but is not linked experimentally to bodily signals, and the rAI, that appears to be involved in the conscious perception of heartbeats.

It would logically follow that the self as expressed in the DN is not related to interoceptive signals. Still, the dichotomous view presented above has to be nuanced by a few experimental findings. First, both the DN and the rAI are found differentially activated in studies targeting the *bodily self* [19]. These studies manipulated body ownership and self-location by creating multisensory conflicts between visual and tactile information, and found a consistent involvement of the right inferior parietal lobule [20,21] and the posterior cingulate cortex (PCC) [22], i.e. two nodes of the DN, but also somatosensory regions and the insular cortex [23]. Second, the meta-analysis of the self cited above [18] focused on midline structures and showed a consistent link between midline nodes of the DN and the self, but did not draw any conclusion on the link between insula and self. Conversely, while the DN is not particularly known for being involved in autonomic regulation, we showed the existence of neural responses to heartbeats in the DN [24], which are markers of the neural processing of ascending cardiac information. We further revealed a direct link between the self and neural responses to heartbeats in the DN [25]. In a whole-brain analysis of magnetoencephalography (MEG) data, we found that the amplitude of neural responses to heartbeats in the two midline nodes of the DN (the PCC and the ventromedial prefrontal cortex, vmPFC) encoded the involvement of the self in spontaneous thoughts. These results suggest that the cardiac monitoring function of the DN is related to the neural implementation of the self.

Here, we hypothesize that a common mechanism, the neural response to heartbeats, could underlie the self in both the medial DN and the rAI. The objectives of this article are threefold. First, in a new meta-analysis of the literature, we confirm the link between the self and DN, and test the link between the self and rAI. We also probe the overlap of DN and rAI with regions involved in autonomic regulation to strengthen our proposal that visceral functions of the DN have been underestimated [25]. Second, we aim at confirming the link between neural responses to heartbeats in the DN and the self with intracranial electroencephalography (iEEG) in epileptic patients. Third, we test whether neural responses to heartbeats in the insula contribute to the self, using both iEEG and a region-of-interest (ROI) approach of the MEG data of healthy participants presented in [25].

Both patients and healthy participants performed a thought-sampling task (figure 1*a*), where they had to fixate a point on the screen and let their mind wander freely for 13–30 s until a visual stimulus was displayed. They had to evaluate the thought they were having at the moment of stimulus display on two scales that targeted two aspects of the self (figure 1*c*). Participants evaluated on the 'I' scale their involvement in the thought as the subject or agent, the one who acts, feels or perceives from the first-person perspective. Ratings on the 'I' scale were high for thoughts such as 'I have to make a phone call' or 'I am thirsty', and low for thoughts with little engagement of the 'I' such as 'It's raining' or 'He is coming tomorrow'. Participants evaluated on another scale to what degree they were thinking about themselves ('Me' scale). Ratings on the 'Me' scale were high when participants were thinking about themselves, such as in 'I am thirsty' or 'I should be more concerned', but low when the thought was directed towards something or someone else, as in 'It's raining' or 'I will make a phone call'. We measured HERs preceding the display of the visual stimulus (figure 1*b*), and correlated the amplitude of HERs during the thought with the ratings on the 'I' and the 'Me' scales.

2. Material and methods

(a) Patients

Five epileptic patients (mean age = 27.6, s.d. = 7.2; two males; right-handed; see the electronic supplementary material, table S1) gave their written informed consent to participate in this study. These patients suffered from drug-refractory focal epilepsy and were implanted stereotactically with depth electrode shafts as part of a presurgical evaluation. Implantation sites were selected on clinical criteria only, without reference to the present protocol. None of the patients had brain lesions, dysplasia nor substantial cognitive impairments. This experiment was approved by the ethics committee of Pitié-Salpêtrière Hospital (Comité de Protection des Personnes).

(b) Intracranial electroencephalography procedure

The thought-sampling paradigm used here corresponds to the one developed by Babo-Rebelo *et al.* [25], where it is explained in full detail. Briefly, patients were presented with three to five blocks of nine trials each (electronic supplementary material, table S1). Each trial consisted of a fixation period followed by a visual stimulus. Fixations ranged from 13.5 to 29.9 s and were randomized in each block. Participants were asked to let their mind wander as naturally as possible during fixation and to press a button in response to the visual stimulus. Then, they rated the thought they were having at the moment of display of the visual stimulus, along four continuous scales. The 'Actor/Author' scale targeted the 'I' dimension of the self ('I' scale) and evaluated the degree to which the participant was seeing or feeling himself/herself as the actor or author during the thought. Participants were instructed to use high ratings ('+') when they were adopting their own perspective, i.e. when they were the protagonist or the agent of the thought, as in 'I will make a phone call'. Low ratings ('−') were used when someone else was the protagonist of the thought ('His office is far away') or nobody in particular ('It's raining'). The 'Content' scale targeted the 'Me' dimension of the self ('Me' scale), i.e. how much the thought was focused on the participant himself/herself or on something external. The 'Me' extreme

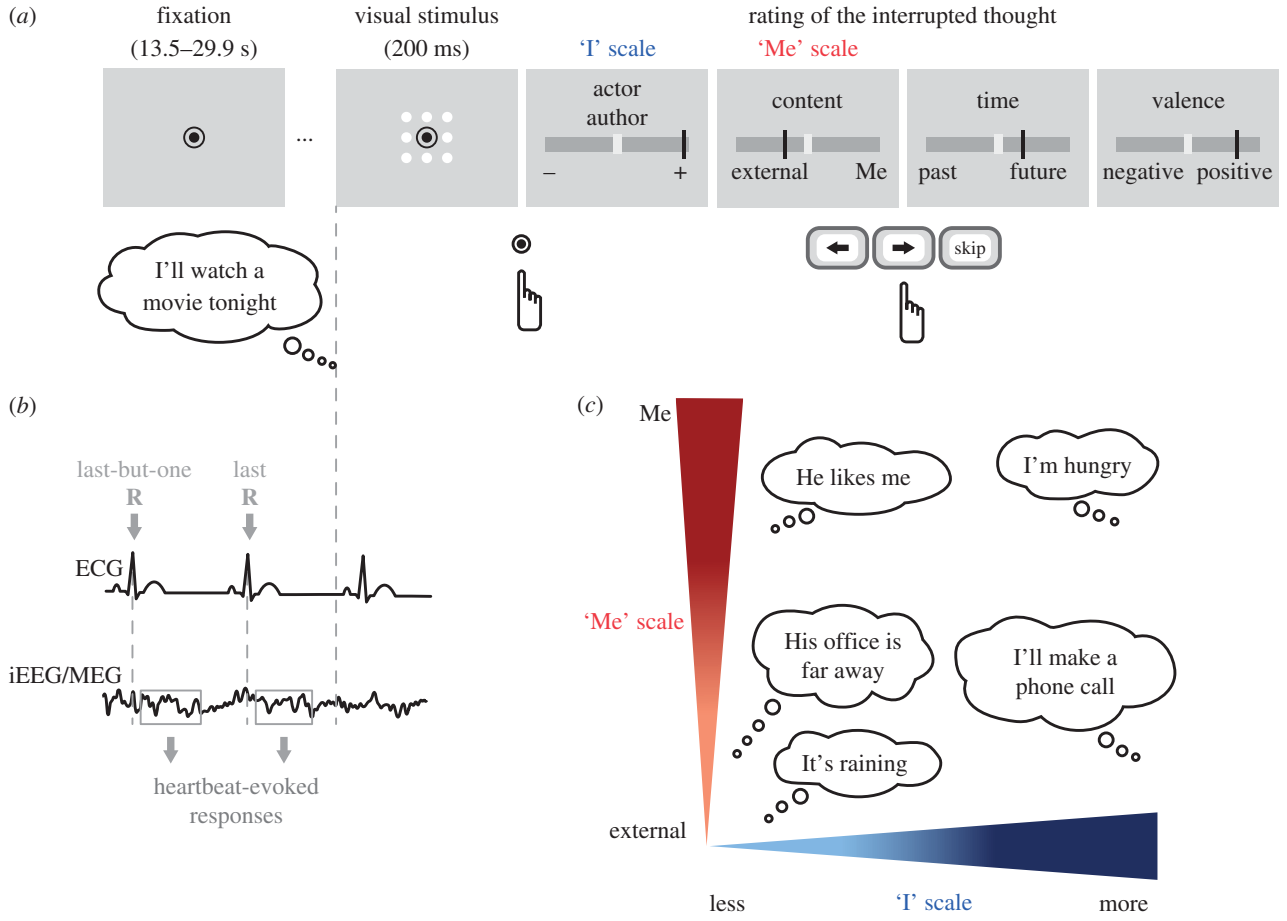


Figure 1. Experimental paradigm. (*a*) Time course of a trial. Each trial consisted of a fixation period interrupted by a visual stimulus. During fixation, participants were asked to let their thoughts develop freely. Participants pressed a button in response to the visual stimulus and had to remember the thought that was interrupted by the visual stimulus. They rated this thought along four scales ('I', 'Me', Time and Valence) or could skip the ratings if the interrupted thought was unclear or if they were not sure how to use the scales. (*b*) Intracranial electroencephalographic (iEEG) and magnetoencephalographic (MEG) data were locked to the two R-peaks of the electrocardiogram (ECG) preceding the visual stimulus, to compute HERs during the thought. (*c*) Examples of thoughts along the two scales of self-relatedness. The 'I' scale described the engagement of the participant as the protagonist or the agent in the thought. The 'Me' scale described the content of the thought, that can be oriented either toward oneself or toward an external object, event or person (adapted from [25]).

of the scale was to be used when participants were thinking about themselves, about their feelings, body or mood, as in 'I'm hungry', 'I should be more concerned' or 'I'm bored'. The 'External' extreme was to be used when participants were thinking about something that was external to them, as for instance 'It's raining' or 'What was the title of the book that Peter recommended?'. The 'Time' scale was used to report whether the thought referred to past, present or future events, while the 'Valence' scale was used to determine whether the thought was pleasant or unpleasant. Participants could skip the ratings if they did not have any clear thought when the stimulus appeared or if they did not know how to rate the thought. If a trial was skipped a new one was added to the block, unbeknownst to the participant.

Before performing the actual experiment, patients were given written and oral instructions and were trained on the scales by rating five examples of thoughts. Their ratings were discussed with the experimenter to ensure task comprehension. Patients then performed a practice block of the thought-sampling task, with two trials.

(c) Intracranial electroencephalographic data acquisition, preprocessing and electrode localization

Patients were implanted intracerebrally with 7–13 depth electrode shafts, each bearing 3–12 contacts (Ad-Tech platinum electrodes with a diameter of 1 mm and 5 mm between contacts). iEEG and electrocardiogram (ECG) data were acquired simultaneously, with either a Micromed (two patients, sampling rate: 1024 Hz; online band-pass filter: 0.15–463.3 Hz; reference: Cz electrode) or Neuralynx monitoring system (three patients; sampling rate: 4000 Hz; online low-pass filter: 1000 Hz; reference: electrode contact in the skull).

Data were downsampled to 1000 Hz and band-pass filtered off-line between 0.5 and 25 Hz, using a fourth-order Butterworth filter. All iEEG signals were re-referenced to their nearest neighbour on the same electrode shaft (bipolar montage) to limit volume-conducted influences, including the cardiac-related artefact. In the following, we will refer to these bipolar montages as 'recording sites'.

Electrode contacts were automatically identified on the computed tomography (CT)-scan obtained after electrode implantation, using a watershed transform-based algorithm. The CT-scan and magnetic resonance imaging (MRI) obtained after implantation were registered to the pre-implantation MRI using Baladin [26], and all images were normalized to Montreal Neurological Institute (MNI) space using SPM12. The automatic electrode localization was verified visually and corrected if necessary using an interactive tool (EpiLoc toolbox developed by the STIM (Stereotaxy: Techniques, Images, Models) engineering platform (http://icm-institute.org/en/cenir-stim-stereotaxy-core-facility-techniques-images-models-2/) in the Institut du Cerveau et de la Moelle Epinière, in Paris). The coordinates of each recording site are reported as the coordinates of the midpoint between the two corresponding contacts.

(d) Rationale for intracranial electroencephalographic analyses

iEEG analyses were restricted to a subsample of recording sites selected on the basis of their distance to the regions where previous MEG results [25] were found (posteromedial cortex, vmPFC) or where we defined *a priori* ROIs (insula). For each region, we defined a volume of interest as the union between the MEG cluster and the corresponding functional territories. For instance in posteromedial cortex, we considered the voxels in the MEG cluster as well as the voxels belonging to the ventral precuneus and ventral posterior cingulate cortex (vPCC). We selected recording sites inside, or at less than 6 mm from the borders of this volume. This limit of 6 mm corresponds to a fair approximation of the borders of these regions, considering both the smoothness applied to functional MRI and MEG source localization masks, but also considering the accuracy of bipolar intracranial recordings.

Even though all patients responded to all scales at each trial, we only analysed the data corresponding to the scale of interest given the MEG results. Therefore, only the 'I' scale was analysed for recording sites in the posteromedial cortex (patient 4) and in the insular region (patients 3 and 5), and only the 'Me' scale was analysed for recording sites in the vmPFC (patients 1, 2 and 3).

(e) Intracranial heartbeat-evoked responses analysis

To detect R-peaks in the ECG, we correlated the *z*-scored ECG signal with a template QRS complex created for each patient and identified the local maxima within episodes of correlation larger than a threshold chosen for each patient. R-peak detection was verified by checking for the absence of outliers in the inter-beat-interval distribution as well as by visual inspection in a time window from -6 to 3 s relative to the visual stimulus.

Epochs of iEEG data were extracted from -100 to 600 ms relative to the two R-peaks preceding each visual stimulus by at least 700 ms. Epochs that exceeded ± 200 mV, which showed a dynamic range of 300 mV or more in a 20 ms interval were excluded from analysis. Data were subsequently visually inspected to discard any additional epochs with excessive noise or epileptic activity. Because recording sites of interest were far from epileptic foci, we discarded only a few epochs (less than 14.8% of the trials in all patients). The final number of trials used in the analysis is reported in the electronic supplementary material, table S1. For each trial, we averaged the two obtained epochs, resulting in one HER per trial and per recording site of interest.

We aimed at testing for each recording site whether the amplitude of HERs was modulated by the self-relatedness of ongoing thoughts. For each time point t of the HER, we computed across trials the Pearson correlation between the *z*-scored HER amplitude at time t and the corresponding *z*-scored rating of the thought on the scale being tested. We then obtained a time course of Pearson correlations and a time course of t-values of the Pearson correlation, revealing the amount of correlation between HER amplitude and ratings on the scale of interest at each time point of the HER. We here used a correlational approach at the single trial level rather than comparing the average HERs for trials rated as high and for trials rated as low as in MEG data [25] to take advantage of the higher signal-to-noise ratio of the iEEG data.

To look for time windows where HER amplitude significantly correlates with ratings, while correcting for multiple comparisons over the time domain, we applied a cluster-based permutation test [27] on the two-tailed t-values of Pearson's correlation across time samples of the time window 300–600 ms relative to the R-peak, for each recording site. Briefly, individual samples with a t-value corresponding to a p-value below an arbitrarily selected threshold ($p < 0.05$, two-tailed) are clustered together based on temporal adjacency. Clusters are characterized by the sum of t-values of the individual samples. To establish the likelihood that a cluster was obtained by chance, we shuffled 10 000 times the ratings with respect to the HERs and repeated the clustering procedure selecting the maximum positive cluster-level statistic and the minimum negative cluster-level statistic. The Monte Carlo p-value corresponds to the proportion of elements in the distribution of maximal (or minimal) cluster-level statistics that exceeds (or is inferior to) the originally observed cluster-level test statistics and is intrinsically corrected for multiple comparisons on time samples. The statistical tests were restricted to the time window 300–600 ms post R-peak and not to the entire HER, because this time window is known to be devoid of the cardiac-field artefact [28]. We also applied a Bonferroni correction on the Monte Carlo p-values, to account for the number of recording sites tested per patient.

Note that here HERs were locked to R-peaks, not to T-peaks as in [25], because T-peaks could not be reliably identified on the ECG signal that had a lower signal-to-noise ratio in clinical settings. To compare latencies between the previous MEG [25] results and the results presented in the current paper both in MEG and iEEG, one has to keep in mind that the average R-T interval in the MEG data is 269 ms.

(f) Surrogate heartbeats

To demonstrate that the observed effects were locked to heartbeats, we checked whether the correlations between HER amplitude and ratings could be obtained with the same sampling of the neural data but unsynchronized with heartbeats. We created 1000 permutations of heartbeats, where the timing of the pair of heartbeats of trial i in the original data is randomly assigned to trial j. The same criteria for rejecting artefactual epochs and computing of HERs was applied. For each permutation, we obtained a set of neural responses to surrogate heartbeats and computed the cluster summed t statistics as described above. For each permutation, we extracted the smallest sum of t-values for recording site 2 of patient 1 and recording site 1 of patient 4 (because the original sum of t-values was negative), and the largest sum of t-values for the recording site 2 of patient 3 (because the original sum of t-values was positive). We then compared the distribution of those surrogate values with the observed original sum of t-values. This control was performed on iEEG data (for MEG data, see [25]).

(g) Region-of-interest analysis on magnetoencephalographic data

We here used an ROI approach centred on the insula, to analyse the MEG data of Babo-Rebelo *et al.* [25]. Sixteen healthy participants (mean age: 24.1 ± 0.6 yr, eight males) performed five blocks of 16 trials of the thought-sampling task, while MEG activity (Elekta Neuromag TRIUX with 102 magnetometers and 204 gradiometers, sampling rate of 1000 Hz, online low-pass filtered at 330 Hz) was acquired simultaneously with ECG activity (seven electrodes around the neck, 0.03–330 Hz). MEG and ECG data were band-pass filtered between 0.5 and 25 Hz. The cardiac-field artefact was corrected on the MEG data using an independent component analysis. HERs were obtained at the sensor level by averaging brain activity locked to the two R-peaks preceding the visual stimulus. For each scale, trials were median split and an average HER was computed for trials rated as 'high' and for trials rated as 'low'. We here used a median split approach because analyses were done at the group level, on data that has a lower signal-to-noise ratio compared with iEEG. Source localization of the HERs was performed with the BrainStorm toolbox [29], using a model consisting of 15 002 current dipoles from the combined time series of magnetometer and gradiometer signals using a linear inverse estimator (weighted minimum-norm current estimate). We created BrainStorm scouts using the niftii masks from Deen *et al.* [30], to

identify the vertices corresponding to the three right insular ROIs: posterior insula (PI, 50 vertices), ventral anterior insula (vAI, 82 vertices) and dorsal anterior insula (102 vertices). Dorsal and vAI scouts had 33 vertices in common, owing to the low resolution of the source model relative to the MRI masks. We then averaged the neural currents corresponding to each of the scouts, and compared, for each ROI, the average cortical current corresponding to 'high' ratings with the one corresponding to 'low' ratings, on the 'I' and 'Me' scales separately. To assess the statistical difference in HERs between 'high' and 'low' ratings, while controlling for multiple comparisons over the time domain, we applied as before a cluster-based permutation test for each ROI, but based on the *t*-test between 'high' and 'low' conditions. The resulting Monte Carlo *p*-values were Bonferroni corrected for testing on two different scales ('I' and 'Me').

We also tested for a correlation between ROI results and individual interoceptive abilities, which were measured in the 16 MEG participants using the heartbeat-counting task [5], over six blocks of variable durations (30–120 s) [25]. Interoceptive abilities were not measured in patients.

(h) Meta-analysis of the 'self'

This meta-analysis was performed using the Neurosynth platform (http://neurosynth.org) [31] that contains nearly 11 400 neuroimaging studies (May 2016). From each article, Neurosynth automatically extracts a set of terms that occur at a high frequency (greater than 1 in 1000 words) and the activation coordinates reported in the study (coordinates are transformed to MNI space if necessary). The database currently contains 3107 terms. We explored the term 'self', which appeared in 903 studies and encompassed 33 560 activations. The automated meta-analysis corresponds to a statistical inference map, from the comparison of coordinates reported in studies containing the term 'self' with coordinates from studies that do not contain the term. The forward inference map corresponds to *z*-scores of the likelihood that a voxel will be activated if a study uses the term 'self' (P(Activation|Term)). The forward inference map thus corresponds to regions that are *consistently* active in studies related to the self, but that may also be active in other paradigms not related to the self. The reverse inference map reports the *z*-scores corresponding to the likelihood that 'self' is used in a study given the presence of reported activation in a particular voxel (P(Term|Activation)). The reverse inference map therefore corresponds to regions that are *selectively* associated with the word 'self'. The reverse inference map controls for base rate differences between regions, so regions that lack selectivity (i.e. regions that are associated with many different terms) are not included in the map. Both maps were corrected for multiple comparisons using a false-discovery rate (FDR) approach, with an FDR of 0.01, meaning that about 1% of activated voxels are false positives, as intrinsically implemented in the Neurosynth platform.

(i) Overlap between our results and anatomical parcellations and meta-analyses

We used a structural MNI152 template image on mricron (https://www.nitrc.org/projects/mricron) to represent our results and the overlap with parcellations and meta-analyses. The MNI coordinates of the vertices showing significant differential HER activity in a previous MEG study [25] were obtained using the BrainStorm functions cs_scs2mri and cs_mri2mni. Niftii masks were then created, displaying the significant voxels, that were expanded (we considered a square of three voxels side, centred on the significant voxel) to facilitate visualization. These masks were then overlaid with the parcellation of the posteromedial cortex from [32], the parcellation of the vmPFC from [33] and the parcellation of the insula from [30]. We also overlaid all results with masks resulting from a meta-analysis on the autonomic brain as described in [34]. All masks were transformed to a final dimension of $91 \times 109 \times 91$, using the function ImCalc of SPM12. The masks of the posteromedial cortex [32], the vmPFC [33], the insula [30] and of the autonomic brain [34] were provided by the corresponding authors.

3. Results

(a) The self and autonomic regulation

To evaluate the contribution of the DN and rAI to the self, as well as their overlap with regions involved in autonomic regulation, we first conducted an automated meta-analysis [31] of 903 studies pertaining to the self. This analysis confirms on a large dataset that the DN is *selectively* related to the self (figure 2; electronic supplementary material, table S2, reverse inference): activity in the DN is likely to indicate self-related processing. The insula is *consistently* activated in the literature related to the self (figure 2; electronic supplementary material, table S3, forward inference), but is not *selective* of the self. In other words, differential activation in the insula can pertain to the self but can also be found in many other cognitive paradigms.

Regions associated with autonomic regulation [34] overlap with self-related regions in the rAI, but also in the DN: the posterior midline node of the DN is associated with parasympathetic regulation, while the frontal midline node of the DN is associated with sympathetic regulation (figure 2).

(b) The ventromedial prefrontal cortex and the 'Me'

In a previous MEG study [25], we found that the amplitude of HERs correlates with the involvement of the 'Me' dimension in the left vmPFC. A further analysis of the MEG cluster showed that it is located mainly in areas 14 m and 32 of the medial frontal cortex (table 1), according to the anatomical parcellation of Neubert *et al.* [33]. Moreover, 41.3% of the MEG cluster overlapped with sympathetic regulation regions (derived from studies on electrodermal activity) [34] which were also mainly located in areas 14 m and 32 (table 1).

To try and replicate the MEG results with intracranial recordings, we selected recording sites inside 14 m or 32 regions or at less than 6 mm from the borders of these regions (electronic supplementary material, table S4). We therefore analysed three recording sites on the left hemisphere from two different patients. We tested each recording site for a trial-by-trial correlation between the amplitude of HERs and the ratings on the 'Me' scale.

Trial-by-trial HERs were obtained by averaging brain activity locked to the two R-peaks preceding each visual stimulus. We computed at each time point the Pearson's correlation across trials between HER amplitude and the rating of the thought on the 'Me' scale. We then used a clustering procedure, which corrects for multiple comparisons over time, to identify, within the time window 300–600 ms after the R-peak, moments where HER amplitude significantly correlated with ratings on the 'Me' scale.

We found that the amplitude of HERs in recording site 2 of patient 1 (MNI coordinates: −14 38 −16, figure 3*d*) significantly correlated with 'Me' ratings (cluster sum(t) = −9547, Monte Carlo $p = 0.046$, Bonferroni-corrected for the two recording sites tested in patient 1), in the time window

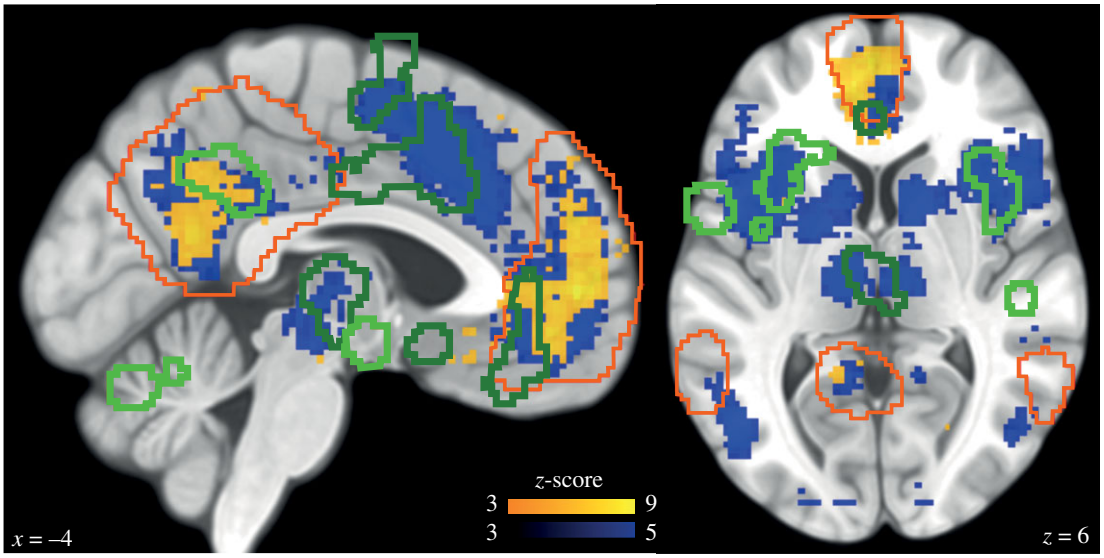


Figure 2. Overlap between DN, self and autonomic regulation meta-analyses. The orange outline represents the DN, as defined in Laird [35]. Green outlines highlight regions responsible for sympathetic (dark green) and parasympathetic (light green) regulation [34]. The results of the automated [31] meta-analysis on the term 'self' are presented in yellow (reverse inference map) and in blue (forward inference map). The sagittal view (left) shows that the reverse inference map of the self is associated with the DN, where it overlaps with autonomic regulation regions. The axial view (right) shows that the rAI is associated with the forward inference map of the self and overlaps with autonomic regulation regions.

Table 1. Percentage distribution of the anterior MEG cluster and of the sympathetic regulation areas [34] on the different sub-regions of the vmPFC [33]. The remaining 29% of the MEG cluster was located in the undetermined territory lying in between those three regions.

	14 m only	32 only	overlap 14 m and 32	11 m
MEG cluster (%)	29	28	4	10
sympathetic regulation areas (%)	46	22	4	5

304–354 ms after the R-peak (mean Pearson correlation coefficient = −0.58; figure 3*a*,*b*). The mean Pearson correlation coefficient in this time window decreased at recording sites that were further away from the midline (figure 3*c*). To show that the observed effects were truly locked to heartbeats and not driven by slow fluctuations of neural activity, we created 1000 permutations of surrogate heartbeats and performed the same analyses on the recording site 2 of patient 1. Only three permutations generated a cluster t statistic exceeding the original one (Monte Carlo $p = 0.003$; electronic supplementary material, figure S2*a*), confirming that these results are indeed locked to heartbeats.

We also tested for a correlation between HER amplitude and 'I' ratings at recording site 2 of patient 1, and found a significant correlation (cluster sum $t = -8424$, Monte Carlo $p = 0.0328$, uncorrected, cluster time window: 306–352 ms after the R-peak). This is different from the group-level analysis of MEG data that revealed a specific effect for the 'Me' in vmPFC [25]. It should be noted that in patient 1, the correlation of the ratings between the 'I' and the 'Me' dimensions was very high (Pearson $r = 0.91$), higher than in other patients (electronic supplementary material, table S1) or healthy participants (electronic supplementary material, table S7). iEEG results in this patient thus confirm that neural responses to heartbeats in vmPFC covary with the self, but do not bring any further information on the dissociation between the 'I' and 'Me' dimensions.

According to the individual anatomy of patient 1 (electronic supplementary material, figure S1*a*), the recording site 2 was located in between the cingulate sulcus, where MEG results were found, and the olfactory sulcus. Recording site 1 of patient 1, that was located more ventrally in the olfactory sulcus (electronic supplementary material, figure S1*a*), did not show any significant correlation (Monte Carlo $p = 0.68$, Bonferroni-corrected for the two recording sites tested in patient 1). In patient 2, a recording site located not in the vicinity of the medial wall but more laterally in the orbitofrontal cortex (fundus of the intermediate orbital sulcus, electronic supplementary material, figure S1*b*) did not show any significant correlation either (no candidate clusters). Altogether, the pattern of results observed with intracranial data is compatible with a neural source in the cingulate sulcus, which is included in the MEG cluster.

MEG results further suggest that HERs in vmPFC are left lateralized. We tested for a null effect at a recording site in the right homologue 14 m region, from a different patient (patient 3). This contact, located in between the olfactory sulcus and the supraorbital sulcus (electronic supplementary material, figure S1*c*), did not show any significant effects (no candidate clusters). This iEEG negative result in the right hemisphere is compatible with the left-lateralization of self-related HERs in vmPFC observed in MEG, but might also be due to an electrode location too ventral to pick activity from the cingulate sulcus and gyrus. Note that a significant effect was observed in this patient at a different location, as described below, indicating that this patient understood the task.

No correlation between heart rate and 'Me' ratings was observed (Pearson correlation between 'Me' ratings and the interval between the two R-peaks preceding the visual stimulus: $r = 0.12$, $t_{25} = 0.58$, $p = 0.57$).

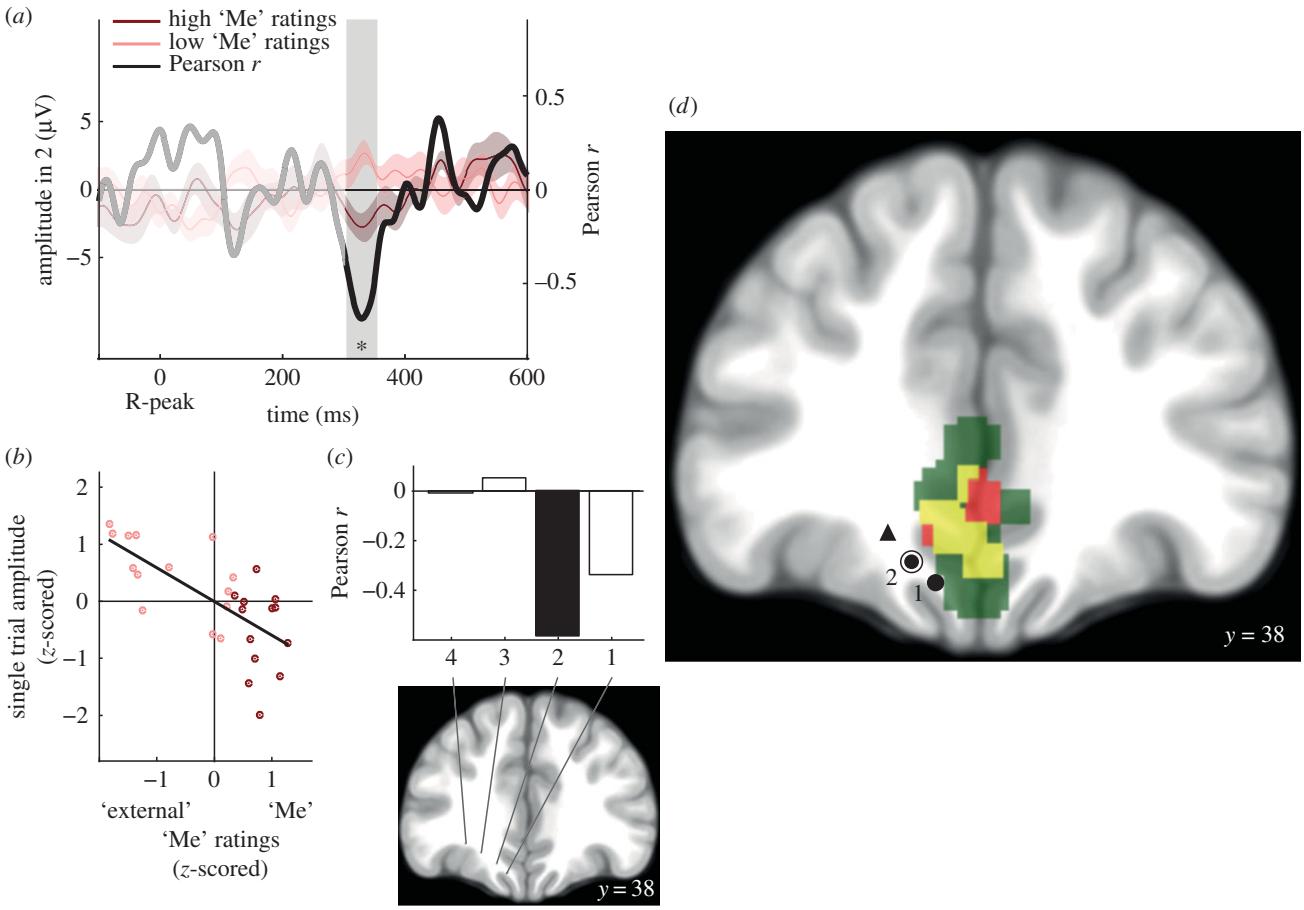


Figure 3. The trial-by-trial amplitude of HERs in the vmPFC correlates with the involvement of the 'Me' in spontaneous thoughts (patient 1). (*a*) Time course of the Pearson correlation coefficient *r* between the trial-by-trial HER amplitude and the ratings on the 'Me' scale (black), and HERs (±s.e.m.) for 'high' (dark red) and 'low' (light pink) ratings on the 'Me' scale (median split of ratings), for recording site 2 (circled dot in (*d*)). The signal that might be residually contaminated by the cardiac-field artefact appears in lighter colour (not included in the analysis). The grey area highlights the time window in which a significant trial-by-trial correlation between HER amplitude and 'Me' ratings was observed. (*b*) HER amplitude in the significant time window plotted against 'Me' ratings. Each point represents one trial. (*c*) Mean Pearson correlation coefficient in the 304–354 ms time window, along the different recording sites of the electrode shaft of patient 1. The black bar corresponds to the recording site for which a significant correlation was found. (*d*) Differential HERs, sympathetic regulation and vmPFC. Recording site 2 (circled dot) showed the significant correlation, while recording site 1 and the triangle (patient 2) showed no effect. Regions in red showed differential responses to heartbeats along the 'Me' scale, in a previous MEG study [25]. Regions in green are involved in sympathetic regulation [34]. Yellow corresponds to the overlap between MEG results and sympathetic regulation regions.

The results of iEEG data thus confirm the existence of HERs distinguishing between different levels of self-relatedness of spontaneous thoughts on the 'Me' scale in the cingulate sulcus, at the border between areas 32 and 14 m as identified with MEG. Regions located more ventrally or more laterally did not show the effect. The areas involved respond to heartbeats and thus seem to be monitoring visceral inputs, but they are also involved in sympathetic regulation.

(c) The posteromedial cortex and the 'I'

In MEG data [25], HER amplitude in the left posteromedial cortex was shown to correlate with the involvement of the 'I' in ongoing spontaneous thoughts. By comparing the MEG cluster with the anatomical parcellation of the posteromedial cortex by Bzdok *et al*. [32], we here show that 50.4% of the MEG cluster was located in the left vPCC and 31.5% in the left ventral precuneus (vPrc) (table 2 and figure 4*d*). Interestingly, none of these regions seems to be involved in parasympathetic regulation (mostly derived from high-frequency heart rate variability [34]), which is exclusively associated with the dorsal PCC (figure 4*d* and table 2).

In order to confirm the involvement of the ventral precuneus territory, we analysed two recording sites of patient 4, which were inside the left vPrc or at less than 6 mm from its borders (electronic supplementary material, table S5). We tested for a trial-by-trial correlation between the amplitude of HERs in these recording sites and the ratings on the 'I' scale, in accordance with the MEG results.

We found that the amplitude of HERs recorded in the most medial recording site that was located inside the vPrc region (figure 4*d*; electronic supplementary material, figure S1*d*, recording site 1, MNI coordinates: −3 −53 49) significantly correlated with 'I' ratings (cluster sum(t) = −8395, Monte Carlo $p = 0.041$, Bonferroni-corrected for the two recording sites tested in patient 4) in the time window 444–500 ms after the R-peak (mean Pearson correlation coefficient = −0.37) (figure 4*a*,*b*). Recording site 2, that was located just outside the vPrc region, did not show a significant correlation with 'I' ratings (Monte Carlo $p = 0.38$, Bonferroni-corrected for the two recording sites tested in patient 4). More generally, the average Pearson correlation coefficient in the 444–500 ms time window decreased as we tested recording sites from the same electrode shaft that

Table 2. Percentage distribution of the posterior MEG cluster and of the parasympathetic regulation areas [34] on the different sub-regions of the posteromedial cortex [32]. The remaining 18% of the MEG cluster were located more posteriorly, in the vicinity of the parieto-occipital sulcus and calcarine fissure.

	precuneus	ventral posterior cingulate cortex	dorsal posterior cingulate cortex	retrosplenial cortex
MEG cluster (%)	32	50	0	0
parasympathetic regulation areas (%)	6	0	92	0

were further away from the midline (figure 4*c*). The test on the 1000 permutations of surrogate heartbeats on recording site 1 confirmed that the effects were truly locked to heartbeats (Monte Carlo $p = 0.011$; electronic supplementary material, figure S2*b*). Additionally, we did not observe a correlation between HER amplitude and 'Me' ratings at recording site 1 (no candidate clusters), nor between the cardiac rhythm and 'I' ratings (Pearson correlation between the interval between the two R-peaks preceding the visual stimulus and 'I' ratings: $r = -0.13$, $t_{43} = -0.83$, $p = 0.41$).

Taken together, the MEG and iEEG results from one patient consistently indicate that two sub-regions of the posteromedial cortex, the vPrc and the vPCC, respond differentially to heartbeats depending on the involvement of the 'I' in thoughts, whereas the adjacent dorsal PCC is involved in cardiac rate regulation but not in encoding self-related information.

(d) The right insula and the 'I'

The insula can be sub-divided in three distinct regions: PI, dorsal anterior (dAI) and vAI [30]. Both right dAI and vAI regions are involved in autonomic regulation (figure 5*a*,*e*,*f*,*g* and table 3), as shown by the meta-analysis by Beissner *et al*. [34]. While parasympathetic regulation is uniquely associated with the dAI, sympathetic regulation is equally associated with dAI and vAI (table 3). Even though the PI is a known visceral centre of the brain, it did not appear to be associated with either sympathetic or parasympathetic regulation (table 3). Here, we test whether HER amplitude covaries with self-relatedness, first by an ROI analysis of the MEG data on healthy participants, and then by analysing three iEEG recording sites in the vicinity of the insula.

(i) Region-of-interest analysis of the insula in magnetoencephalographic data of healthy participants

From MEG data obtained in 16 healthy participants, we computed R-locked HERs and the corresponding sources for trials rated as 'high' and trials rated as 'low' (median split of the trials, electronic supplementary material, table S7) on each self-related scale. We then averaged the resulting neural currents for the vertices belonging to each sub-region of the insula, the PI, the dAI and the vAI, according to the parcellation of Deen *et al*. [30] (figure 5*a*). For each sub-region, we searched for time windows where HERs significantly differed between trials rated as 'high' and trials rated as 'low', separately on the 'I' and the 'Me' scales, using a cluster-based permutation *t*-test over the time window 300–600 ms post R-peak.

We found that neural responses to heartbeats in the dorsal and ventral rAI (figure 5*b*,*c*) significantly differed for trials rated as 'high' and trials rated as 'low' on the 'I' scale (dAI: cluster sum(t) $= -296$, Monte Carlo $p = 8 \times 10^{-4}$; vAI: cluster sum(t) $= -283$, Monte Carlo $p = 0.0012$, Bonferroni-corrected for the two scales tested), in the same time window relative to the R-peak (dAI: 384–486 ms; vAI: 384–480 ms). No differences were observed in the PI (figure 5*d*, no candidate clusters) nor for the 'Me' scale in any of the three right insular regions (dAI: Monte Carlo $p = 0.13$; vAI: Monte Carlo $p = 0.33$; PI: Monte Carlo $p = 0.16$, Bonferroni-corrected for the two scales tested). This ROI-based approach in MEG sources thus revealed differential neural responses to heartbeats in the rAI depending on the involvement of the 'I' in thoughts. The map of the *t*-values associated with the ROI effect (figure 5*e*) shows that there are two foci contributing to the rAI effect, one more posterior and another one more anterior, extending outside the rAI into the inferior frontal gyrus.

We then tested the lateralization of this result, by probing the left dorsal and left vAI. No significant differences between 'high' and 'low' ratings on the 'I' scale were observed (all Monte Carlo $p > 0.3$). In addition, an ANOVA on brain currents averaged over the time window of the significant difference, with hemisphere (left and right) and condition ('high' and 'low') as factors revealed an interaction between hemisphere and condition in both dAI and vAI (dAI: interaction: $F_{1,15} = 7.67$, $p = 0.014$, main effects: $p > 0.14$; vAI: interaction $F_{1,15} = 7.73$, $p = 0.014$, main effect hemisphere: $F_{1,15} = 4.72$, $p = 0.046$, main effect condition: $F_{1,15} = 2.04$, $p = 0.17$). The amplitude of the effects was not modulated by individual interoceptive abilities (Pearson correlation between the difference in HER amplitude between 'high' and 'low' 'I' ratings and interoceptive scores, dAI: mean $r = 0.08$, $t_{14} = 0.3$, $p = 0.8$; vAI: mean $r = -0.03$, $t_{14} = -0.1$, $p = 0.9$).

(ii) Intracranial electroencephalographic analysis of three recording sites in the vicinity of the insula

We then analysed the iEEG data from two patients (3 and 5) who had recording sites at less than 6 mm of the borders of the right dAI (electronic supplementary material table S6). Because MEG results indicated a link between rAI and the 'I' scale, we searched for a trial-by-trial correlation between the HER amplitude at these recording sites and the ratings on the 'I' scale. The clustering test revealed a significant correlation between HER amplitude and 'I' ratings, at the most dorsal recording site (recording site 2 in patient 3; figure 5*f*), at a latency of 397–443 ms after the R-peak (figure 5*g*,*h*; Pearson correlation coefficient $= 0.46$, cluster sum(t) $= 5609$, Monte Carlo $p = 0.014$, Bonferroni-corrected for the two sites tested in this patient). The significant time window in iEEG data from this site is included in the time window where significant effects are found in MEG data. Moreover, the mean Pearson

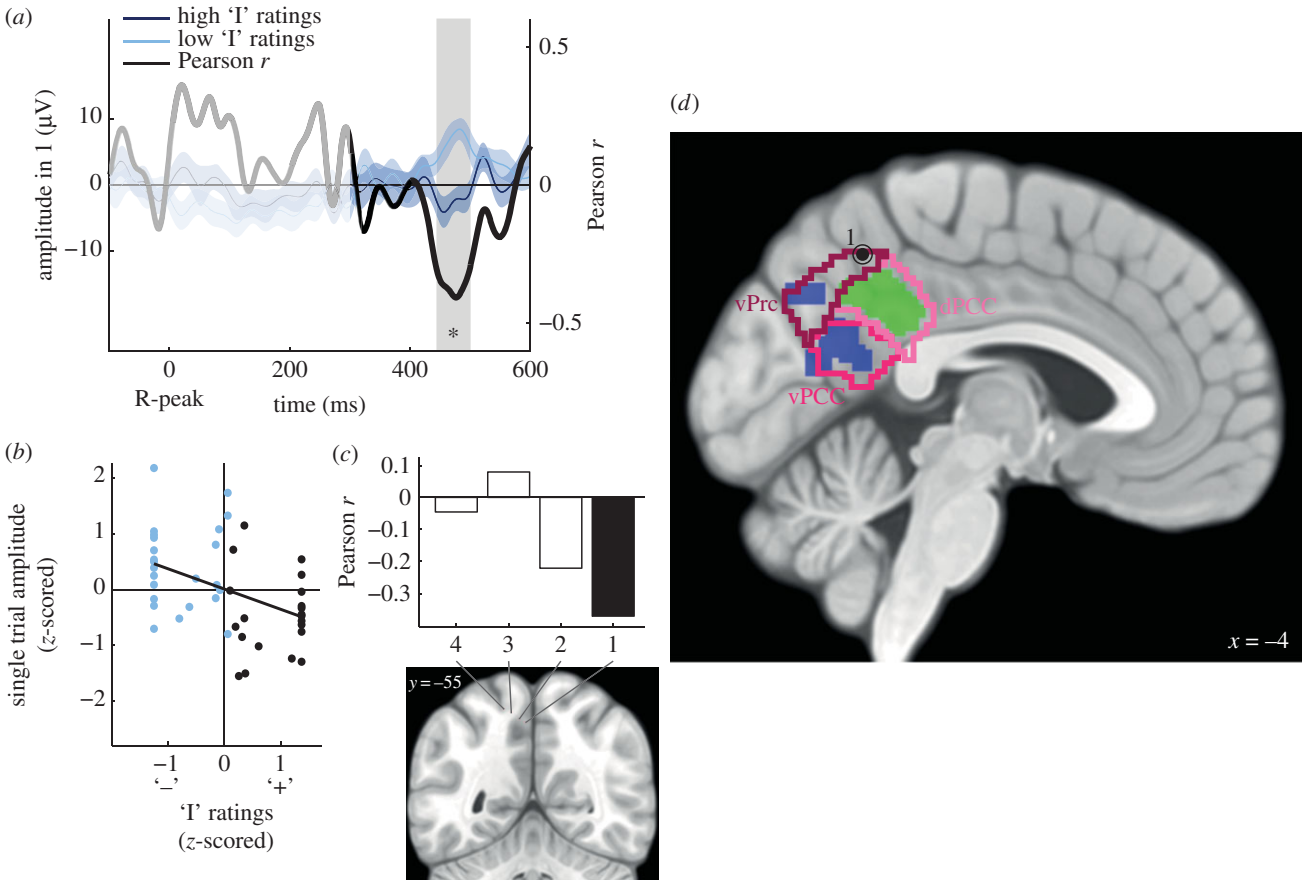


Figure 4. The trial-by-trial amplitude of HERs in the ventral precuneus and vPCC correlates with the involvement of the 'I' in spontaneous thoughts (patient 4). (*a*) Time course of the Pearson correlation coefficient *r* between the trial-by-trial HER amplitude and the ratings on the 'I' scale (black), and HERs (±s.e.m.) for 'high' (dark blue) and 'low' (light blue) ratings on the 'I' scale (median split of ratings), for recording site 1. The signal that might be residually contaminated by the cardiac-field artefact appears in lighter colour (not included in the analysis). The grey area highlights the time window in which a significant trial-by-trial correlation between HER amplitude and 'I' ratings was observed. (*b*) HER amplitude in the significant time window plotted against 'I' rating. Each point represents one trial. (*c*) Mean Pearson correlation coefficient in the 444–500 ms time window, along the different recording sites of the electrode shaft. The black bar corresponds to the recording site for which a significant correlation was found. (*d*) Differential HERs, parasympathetic regulation and posteromedial cortex. The circled dot indicates the location of recording site 1. Regions in blue showed differential responses to heartbeats along the 'I' scale, in a previous MEG study [25]. Regions in green are involved in parasympathetic regulation [34]. Outlines correspond to the parcellation of the posteromedial cortex [32]: ventral precuneus (vPrc, dark pink), dorsal cingulate (dPCC, light pink) and ventral cingulate cortex (vPCC, pink).

correlation coefficient in this time window decreased for recording sites that were located further away from the insular cortex (figure 5*i*). The result was truly locked to heartbeats (Monte Carlo $p = 0.001$, electronic supplementary material, figure S2*c*). Additionally, HER amplitude did not correlate with 'Me' ratings, for recording site 2 of patient 3 (Monte Carlo $p = 0.48$, uncorrected). Last, these results were not associated with a correlation between heart rate and 'I' ratings (Pearson correlation between 'I' ratings and the interval between the two R-peaks preceding the visual stimulus: $r = -0.05$, $t_{24} = -0.25$, $p = 0.80$).

The recording site where we found a significant effect was located at the anterior and dorsal border of the dAI (figure 5*f*; electronic supplementary material figure S1*e*). The other recording site in the same patient was located more ventrally and did not show any significant effect (recording site 1: Monte Carlo $p = 0.53$; Bonferroni-corrected for the two recording sites tested in patient 3). The recording site of patient 5 was located even more ventrally and did not display any significant correlation (Monte Carlo $p = 0.50$ figure 5*j*; electronic supplementary material, figure 1*f*).

iEEG data thus only partially confirm MEG results, with positive results in one patient out of two. Still, the pattern of results in both MEG and iEEG indicate that at least in its most anterior and dorsal part, the rAI generates HERs, the amplitude of which depends on the involvement of the 'I' in spontaneous thoughts.

(e) Comparison of magnetoencephalographic results across vPrc/vPCC, vmPFC and rAI

Here, we used an ROI analysis to show that the rAI is differently responding to heartbeats depending on the self-relatedness of thoughts. However, the rAI did not appear in the whole-brain analysis, as opposed to the midline regions of the DN, the vPrc/vPCC and the vmPFC. We thus attempted at characterizing further the effects in rAI, to understand why this effect was not present in the whole-brain analysis.

We first looked at effect sizes (figure 6). We averaged source amplitudes separately in the vmPFC and vPrc/vPCC clusters derived from the whole-brain analysis, and in the rAI region, defined as the union of dAI and vAI that both showed an effect in the ROI-based approach. Effect size is 3.6 times smaller in rAI than in the vPrc/vPCC and five times smaller than in the vmPFC. Effect size comparison remains difficult to interpret because voxels were selected on the basis of a

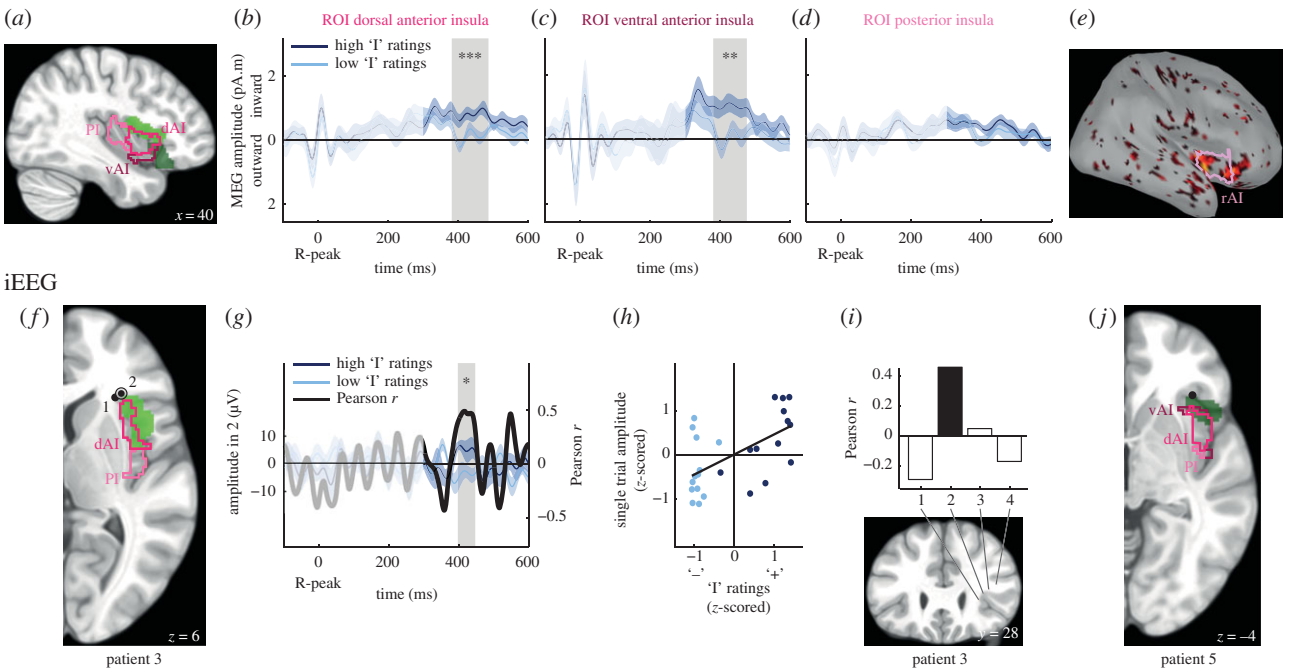


Figure 5. The amplitude of HERs in the rAI correlates with the involvement of the 'I' in spontaneous thoughts, in MEG (panels *a*–*e*) and in iEEG (panels *f*–*j*). (*a*) Sagittal view of the insula with three insular sub-regions [30] highlighted: PI (light pink), dAI (pink), vAI (dark pink). Light green and dark green regions are associated with parasympathetic and sympathetic regulation respectively [34]. (*b*–*d*) Time course of the HER (±s.e.m. across the 16 participants) for 'high' and 'low' responses on the 'I' scale (median split of responses), for the three ROIs in MEG source analysis. The grey area highlights the time window where a significant difference between HERs for 'high' and 'low' ratings on the 'I' scale was observed. (*e*) Differential MEG source activity for 'high' versus 'low' ratings on the 'I' scale averaged over 384–480 ms post R-peak (threshold for visualization: uncorrected $p < 0.05$; 75% smoothness applied to the cortical surface). The pink region corresponds to the rAI (union of dAI and vAI). (*f*) Axial view of the right hemisphere showing the PI (light pink) and dAI (pink). Black dots correspond to the two recording sites analysed in patient 3. Recording site 2 (circled dot) showed a significant correlation between HER amplitude and 'I' ratings, in a time window consistent with the MEG results. Areas in green are involved in parasympathetic regulation [34]. (*g*) Time course of the trial-by-trial Pearson correlation coefficient *r* between HER amplitude and 'I' ratings (black), and HERs (±s.e.m) for 'high' (dark blue) and 'low' (light blue) ratings on the 'I' scale (median split of responses), for recording site 2 of patient 3 (circled dot in *f*). The grey area highlights the time window in which a significant correlation between HER amplitude and 'I' ratings was observed. (*h*) HER amplitude in the significant time window plotted against 'I' ratings. Each point represents one trial. (*i*) Mean Pearson correlation coefficient in the 397–443 ms time window, along the different recording sites of the electrode shaft of patient 3. (*j*) Axial view of the right hemisphere showing the PI (light pink) and dAI (pink), for patient 5. The black dot corresponds to the recording site analysed for this patient, where no significant correlation was observed.

statistical threshold in vmPFC and vPrc/vPCC, while voxels in the AI were selected based on an anatomically defined ROI, which includes non-responsive regions (figure 5*e*). We thus compared source amplitude at vertices thresholded at first-level $p < 0.01$ in the rAI, vPrC/vPCC and vmPFC. Effect size remained 1.7 times smaller in rAI than in the vPrC/vPCC and 2.4 times smaller than in vmPFC.

Another reason why the rAI effect was not picked up in the whole-brain analysis is that the clustering procedure employed favours spatial contiguity. As shown in figure 5*e*, it seems that there are two separate sub-regions of the rAI responding differentially to heartbeats, one in the posterior part of the rAI, another one in the anterior part, extending anteriorly in the inferior frontal gyrus.

Overall, our results indicate that the regions showing the most consistent modulation of HER amplitude in relation to the self are the midline nodes of the DN. The rAI appears to be also involved, but to a lesser extent.

4. Discussion

We aimed at confirming and specifying the existence of visceral monitoring functions in the DN and their links with the self, and at testing whether this mechanism could also be at play in the rAI. We first showed that both the DN and the rAI include regions involved in autonomic functions [34]. We confirm the link between the DN and self [18] and show further that the DN is specific to the self, as opposed to the rAI that is associated with the self, but also with many other, non-self-related paradigms. We found that in two patients the trial-by-trial amplitude fluctuations of intracranially recorded HERs in the DN covaried with the trial-by-trial measure of the involvement of the self in spontaneous thoughts, confirming and refining previous MEG results [25]. An ROI approach of the rAI revealed that both in MEG data of healthy participants and in intracranial recordings of one patient out of two, neural responses to heartbeats covaried with the 'I' dimension of the self. None of these results were associated with changes in heart rate.

5. Methodological considerations and limitations

In this study, we combine data from different sources. The MEG source localization results obtained in a group of healthy participants might be spatially inaccurate, but participants could be trained and task comprehension could be tested and quantified. iEEG data have high spatial accuracy

Table 3. Percentage distribution of insular sympathetic and parasympathetic regulation areas [34] on the different sub-regions of the insular cortex [30]. The sympathetic and parasympathetic insular regions extended over a larger area than the insular parcellation [30].

	dorsal anterior insula	ventral anterior insula	posterior insula
parasympathetic regulation (%)	29	0	1
sympathetic regulation (%)	8	8	0

and good signal-to-noise ratio, but are obtained in patients. In patients, task comprehension was not tested beyond verbal exchanges with the experimenter. Patient 1 for instance seemed not to discriminate between the 'I' and 'Me' dimensions. Because electrode implantation sites are chosen based solely on clinical criteria, electrode coverage of the DN and rAI was not optimal, and only one recording site with positive results could be obtained in each of the three regions explored. Last, all recording sites tested were away from the epileptogenic regions, and did not include epileptic spikes, but more subtle signs of epileptic activity might have gone unnoticed. Despite these pitfalls, there is an overall good agreement between the MEG and iEEG data, as discussed further below, which suggests that MEG localization was rather accurate, and that epileptic patients performed the task in a similar manner as healthy participants.

Another caveat when working on HERs is that cardiac activity can generate two types of artefacts. The cardiac artefact corresponds to the contamination of neural data by the electrical signal of the heart. We analysed time windows that are devoid of this artefact [28] for both MEG and iEEG data, and further corrected MEG data using independent component analysis. The cardiac artefact appeared well suppressed from iEEG data once bipolar derivations are computed (see shaded areas in figures 3*a*, 4*a* and 5*g*). iEEG data are also susceptible to the pulse-related artefact [36] that appears as a slow frequency sinewave or sawtooth pattern. Given the transient nature of the effects reported here in iEEG data, as well as the good agreement between MEG and iEEG latencies and effect durations, it seems unlikely that the pulse-related artefact contributed to the iEEG results.

We also compared the electrophysiological results obtained with iEEG and MEG with MRI results from the literature. MEG source localization is performed on the grey matter ribbon, as can be seen in figure 5*e* and is expressed as a surface. MRI parcellations and functional regions involved in autonomous regulations are expressed in volumes. The conversion between volumes and surfaces might have generated some spatial noise.

6. Heartbeat evoked responses in the default network encode self-relatedness

As in healthy participants, iEEG recordings in epileptic patients show that HERs in the two midline nodes of the DN encoded self-relatedness. Intracranial data in single patients thus confirm the group-level source localization of MEG data in healthy participants [25]. Note that iEEG data confirm that neural responses to heartbeats in vPrC/vPCC are specific to the agentive 'I', but the high correlation between 'I' and 'Me' ratings of the patient implanted in vmPFC does not allow us to tease apart the two dimensions of the self in vmPFC. iEEG data also confirm the temporal order of the effects, with the effect in vmPFC appearing before the effect in vPrc/vPCC. iEEG data further extend the link between HERs and self-relatedness ratings down to the level of single trials, with significant correlations between trial-by-trial HER amplitude and self-relatedness of thought.

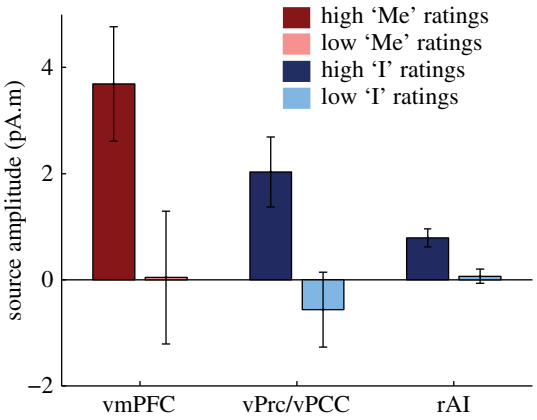


Figure 6. Comparison of effects in the insula and in the DN. Source activity was averaged for each significant time window, across the significant vertices (vmPFC: left, vPrc/vPCC: middle) or across the vertices belonging to the rAI ROI (right).

The detailed anatomical analysis of both iEEG and MEG source-localized results indicates that in vmPFC, the most active regions are areas 14 m and 32 [33], in the ventral part of the anterior cingulate cortex. iEEG recording sites located more laterally or more ventrally did not show any significant effect. Areas 14 m and 32 also contribute to sympathetic regulation [34]. In the posteromedial cortex, HERs varying with self-relatedness occurred in the vPrc and vPCC [32], that are not involved in autonomic regulation, as opposed to the area lying just anterior to them, that is associated with parasympathetic regulation. This result shows that regions that are not associated with autonomic regulation can nevertheless receive and differentially respond to cardiac information, depending on self-relatedness.

Our results thus confirm that the link between the self and DN [18] is expressed in neural responses to heartbeats, and directly support theories grounding the self in the monitoring of internal signals [1–3].

7. Heartbeat evoked responses in the right anterior insula contribute to encoding the 'I'

Although a whole-brain analysis of MEG data revealed significant results only in the DN, a targeted ROI approach of the three sub-divisions of the insula revealed that neural responses to heartbeats in both the dorsal and ventral rAI vary according to the involvement of the 'I' in spontaneous thoughts, around 400 ms after the R-peak. Note that the effect was smaller in

the rAI than in the DN, and appeared to stem from two distinct foci, which may explain why it was not detected in the whole-brain analysis despite a similar sensitivity of MEG to midline DN nodes and insula [37]. Intracranial recordings targeted the most anterior focus of the rAI in two patients. The effect could be detected in one patient out of two only.

These differential responses to heartbeats occurred during a resting state, without any explicit interoceptive task, because participants are not asked to orient attention to their heartbeats. We thus do not know whether these self-related neural responses to heartbeats in the rAI are linked to the modulation of neural responses to heartbeats in explicit interoceptive tasks. The location of the posterior focus in the rAI where we find differential HERs (figure 5*e*) is compatible with the meta-analysis of interoceptive tasks reported in this issue [38]. Still, it should be noted that explicit interoceptive tasks are likely to tap more onto the 'Me' dimension of the self (thinking about oneself) than about the agentive 'I' dimension of the self that we find to be encoded by HERs in the rAI both in MEG and iEEG data. Because rAI is involved in many cognitive studies [10,11] and is not specific to the self, as shown in the meta-analysis presented in figure 2, understanding the contribution of rAI to the self will certainly require further investigations [39].

8. Interplay between vmPFC, rAI, vPrc/vPCC and autonomic control regions

It has sometimes been proposed that the anterior insula is the cortical interoceptive hub, distributing interoceptive information to other cortical areas [3,40]. Our results rather speak in favour of multiple ascending pathways, as described in Critchley [41], and show a stronger effect in the DN than in the rAI. The earliest effects are observed around 400 ms after R peak, in overlapping time windows, i.e. almost simultaneously in the vmPFC and rAI, where the 'Me' and the 'I' self-dimensions are, respectively, encoded. This is in line with known direct projections from subcortical visceral relays to both insula and ventral cingulate regions [42]. The effect in the vPrc/vPCC corresponds to the same self-dimension as in the rAI, but appears later, around 580 ms after R peak, and is more robust. Because vPrc is connected to rAI [32], the weak rAI effect might fuel the more robust vPrc differential response. Alternatively, the vPrc/vPCC effect might be mediated through vmPFC, because the two structures are strongly functionally coupled [32]. In this case, it remains to be explained how the same cardiac inputs can give rise to the encoding of two different dimensions of the self in vPrc/vPCC and vmPFC.

In addition, our results suggest that self-related HERs correspond to a neural monitoring of cardiac information that does not directly translate into cardiac regulation. Indeed, there was no cardiac rhythm changes associated with the effects reported here. In addition, HER locations, that reflect the central monitoring of ascending cardiac information, do not map perfectly on regions involved in autonomic control, that reflect descending regulatory influences. In particular in the posterior medial cortex, the cortical territory involved in high-frequency cardiac regulation is distinct from the two adjacent regions, the ventral precuneus and vPCC, that show self-related HERs. It might be that the cortical cardiac monitoring function, initially devoted to autonomic regulation, has further evolved into a partially distinct process related to selfhood.

9. Conclusion

We here show that the amplitude of neural responses to heartbeats covaries with the self in both the DN and the rAI, although effects are weaker in the rAI. This implies that the literature on the self and DN should consider neural responses to heartbeats, and that conversely the literature relating interoception and the self in the rAI should consider the DN: both structures are related to the self through the same underlying mechanism.

Ethics. All procedures, both in normal participants and in epileptic patients, were approved by the ethics committee of Pitié-Salpêtrière Hospital (Comité de Protection des Personnes). Both normal participants and epileptic patients gave informed consent prior to the experiments.

Data accessibility. The anonymized MEG and iEEG data presented in the article can be obtained upon request to the first and last authors.

Authors' contributions. M.B.-R. and C.T.-B. designed the experiment. M.B.-R. acquired the data. M.B.-R., N.W. and C.T.-B. analysed the data. C.A. and D.H. provided clinical and anatomical expertise. M.B.-R. and C.T.-B. wrote the paper.

Competing interest. We have no competing interest.

Funding. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no 670325) to C.T.-B., as well as by ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*. M.B.-R. was supported by a grant from the Fundação para a Ciência e Tecnologia (SFRH/BD/85127/2012).

Acknowledgements. We thank Imen El Karoui, Katia Lehongre, Vincent Navarro, Virginie Lambecq and paramedical staff for help with epileptic patient data acquisition and preprocessing. We thank Benjamin Deen, Kevin Pelphrey, Florian Beissner, Simon Eickhoff, Danilo Bzdok and Matthew Rushworth for sharing their anatomical and functional data.

References

1. Damasio AR. 1999 *The feeling of what happens: body, emotion and the making of consciousness*. New York, NY: Harcourt.
2. Park H-D, Tallon-Baudry C. 2014 The neural subjective frame: from bodily signals to perceptual consciousness. *Phil. Trans. R. Soc. B* **369**, 20130208. (doi:10.1098/rstb.2013.0208)
3. Craig AD. 2009 How do you feel - now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* **10**, 59–70. (doi:10.1038/nrn2555)
4. Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ. 2004 Neural systems supporting interoceptive awareness. *Nat. Neurosci.* **7**, 189–195. (doi:10.1038/nn1176)
5. Schandry R. 1981 Heart beat perception and emotional experience. *Psychophysiology* **18**, 483–488. (doi:10.1111/j.1469-8986.1981.tb02486.x)
6. Pollatos O, Kirsch W, Schandry R. 2005 Brain structures involved in interoceptive awareness and cardioafferent signal processing: a dipole source

localization study. *Hum. Brain Mapp*. **26**, 54–64. (doi:10.1002/hbm.20121)

7. Canales-Johnson A *et al.* 2015 Auditory feedback differentially modulates behavioral and neural markers of objective and subjective performance when tapping to your heartbeat. *Cereb. Cortex* **25**, 4490–4503. (doi:10.1093/cercor/bhv076)
8. Khalsa SS, Rudrauf D, Feinstein JS, Tranel D. 2009 The pathways of interoceptive awareness. *Nat. Neurosci*. **12**, 1494–1496. (doi:10.1038/nn.2411)
9. Pollatos O, Schandry R, Auer DP, Kaufmann C. 2007 Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Res*. **1141**, 178–187. (doi:10.1016/j.brainres.2007.01.026)
10. Chang LJ, Yarkoni T, Khaw MW, Sanfey AG. 2013 Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cereb. Cortex* **23**, 739–749. (doi:10.1093/cercor/bhs065)
11. Duncan J, Owen AM. 2000 Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci*. **23**, 475–483. (doi:10.1016/S0166-2236(00)01633-7)
12. Menon V, Uddin LQ. 2010 Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct*. **214**, 655–667. (doi:10.1007/s00429-010-0262-0)
13. Buckner RL, Andrews-Hanna JR, Schacter DL. 2008 The brain's default network: anatomy, function, and relevance to disease. *Ann. N.Y. Acad. Sci*. **1124**, 1–38. (doi:10.1196/annals.1440.011)
14. Jerbi K *et al.* 2010 Exploring the electrophysiological correlates of the default-mode network with intracerebral EEG. *Front. Syst. Neurosci*. **4**, 27. (doi:10.3389/fnsys.2010.00027)
15. Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL. 2010 Functional-anatomic fractionation of the brain's default network. *Neuron* **65**, 550–562. (doi:10.1016/j.neuron.2010.02.005)
16. Andrews-Hanna JR, Smallwood J, Spreng RN. 2014 The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N.Y. Acad. Sci*. **1316**, 29–52. (doi:10.1111/nyas.12360)
17. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005 The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl Acad. Sci. USA* **102**, 9673–9678. (doi:10.1073/pnas.0504136102)
18. Qin P, Northoff G. 2011 How is our self related to midline regions and the default-mode network? *Neuroimage* **57**, 1221–1233. (doi:10.1016/j.neuroimage.2011.05.028)
19. Blanke O. 2012 Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci*. **13**, 556–571. (doi:10.1038/nrn3292)
20. Blanke O, Ortigue S, Landis T, Seeck M. 2002 Stimulating illusory own-body perceptions. *Nature* **419**, 269–270. (doi:10.1038/419269a)
21. Ionta S, Heydrich L, Lenggenhager B, Mouthon M, Fornari E, Chapuis D, Gassert R, Blanke O. 2011 Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron* **70**, 363–374. (doi:10.1016/j.neuron.2011.03.009)
22. Guterstam A, Bjornsdotter M, Gentile G, Ehrsson HH. 2015 Posterior cingulate cortex integrates the senses of self-location and body ownership. *Curr. Biol*. **25**, 1416–1425. (doi:10.1016/j.cub.2015.03.059)
23. Tsakiris M, Hesse MD, Boy C, Haggard P, Fink GR. 2007 Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb. Cortex* **17**, 2235–2244. (doi:10.1093/cercor/bhl131)
24. Park H-D, Correia S, Ducorps A, Tallon-Baudry C. 2014 Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nat. Neurosci*. **17**, 612–618. (doi:10.1038/nn.3671)
25. Babo-Rebelo M, Richter C, Tallon-Baudry C. 2016 Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. *J. Neurosci*. **36**, 7829–7840. (doi:10.1523/JNEUROSCI.0262-16.2016)
26. Bardinet E *et al.* 2009 A three-dimensional histological atlas of the human basal ganglia. II. Atlas deformation strategy and evaluation in deep brain stimulation for Parkinson disease. *J. Neurosurg*. **110**, 208–219. (doi:10.3171/2008.3.17469)
27. Maris E, Oostenveld R. 2007 Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190. (doi:10.1016/j.jneumeth.2007.03.024)
28. Dirlich G, Dietl T, Vogl L, Strian F. 1998 Topography and morphology of heart action-related EEG potentials. *Electroencephalogr. Clin. Neurophysiol*. **108**, 299–305.
29. Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM. 2011 Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci*. **2011**, 879716. (doi:10.1155/2011/879716)
30. Deen B, Pitskel NB, Pelphrey KA. 2011 Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* **21**, 1498–1506. (doi:10.1093/cercor/bhq186)
31. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011 Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670. (doi:10.1038/nmeth.1635)
32. Bzdok D, Heeger A, Langner R, Laird AR, Fox PT, Palomero-Gallagher N, Vogt BA, Zilles K, Eickhoff SB. 2014 Subspecialization in the human posterior medial cortex. *Neuroimage* **106**, 55–71. (doi:10.1016/j.neuroimage.2014.11.009)
33. Neubert F-X, Mars RB, Sallet J, Rushworth MFS. 2015 Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc. Natl Acad. Sci. USA* **112**, E2695–E2704. (doi:10.1073/pnas.1410767112)
34. Beissner F, Meissner K, Bär K-J, Napadow V. 2013 The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. *J. Neurosci*. **33**, 10 503–10 511. (doi:10.1523/JNEUROSCI.1103-13.2013)
35. Laird AR *et al.* 2011 Behavioral interpretations of intrinsic connectivity networks. *J. Cogn. Neurosci*. **23**, 4022–4037. (doi:10.1162/jocn_a_00077)
36. Kern M, Aertsen A, Schulze-Bonhage A, Ball T. 2013 Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. *Neuroimage* **81**, 178–190. (doi:10.1016/j.neuroimage.2013.05.042)
37. Ahlfors SP, Han J, Belliveau JW, Hämäläinen MS. 2010 Sensitivity of MEG and EEG to source orientation. *Brain Topogr*. **23**, 227–232. (doi:10.1007/s10548-010-0154-x)
38. Schulz SM. 2016 Neural correlates of heart-focused interoception: a functional magnetic rosonance imaging meta-analysis. *Phil. Trans. R. Soc. B* **371**, 20160018. (doi:10.1098/rstb.2016.0018)
39. Salomon R, Ronchi R, Donz J, Bello-Ruiz J, Herbelin B, Martet R, Faivre N, Schaller K, Blanke O. 2016 The insula mediates access to awareness of visual stimuli presented synchronously to the heartbeat. *J. Neurosci*. **36**, 5115–5127. (doi:10.1523/JNEUROSCI.4262-15.2016)
40. Damasio A, Carvalho GB. 2013 The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci*. **14**, 143–152. (doi:10.1038/nrn3403)
41. Critchley HD, Harrison NA. 2013 Visceral influences on brain and behavior. *Neuron* **77**, 624–638. (doi:10.1016/j.neuron.2013.02.008)
42. Vogt BA, Derbyshire SWG. 2009 Visceral circuits and cingulate-mediated autonomic functions. In *Cingulate neurobiology and disease* (ed. BA Vogt), pp. 219–236. Oxford, UK: Oxford University Press.

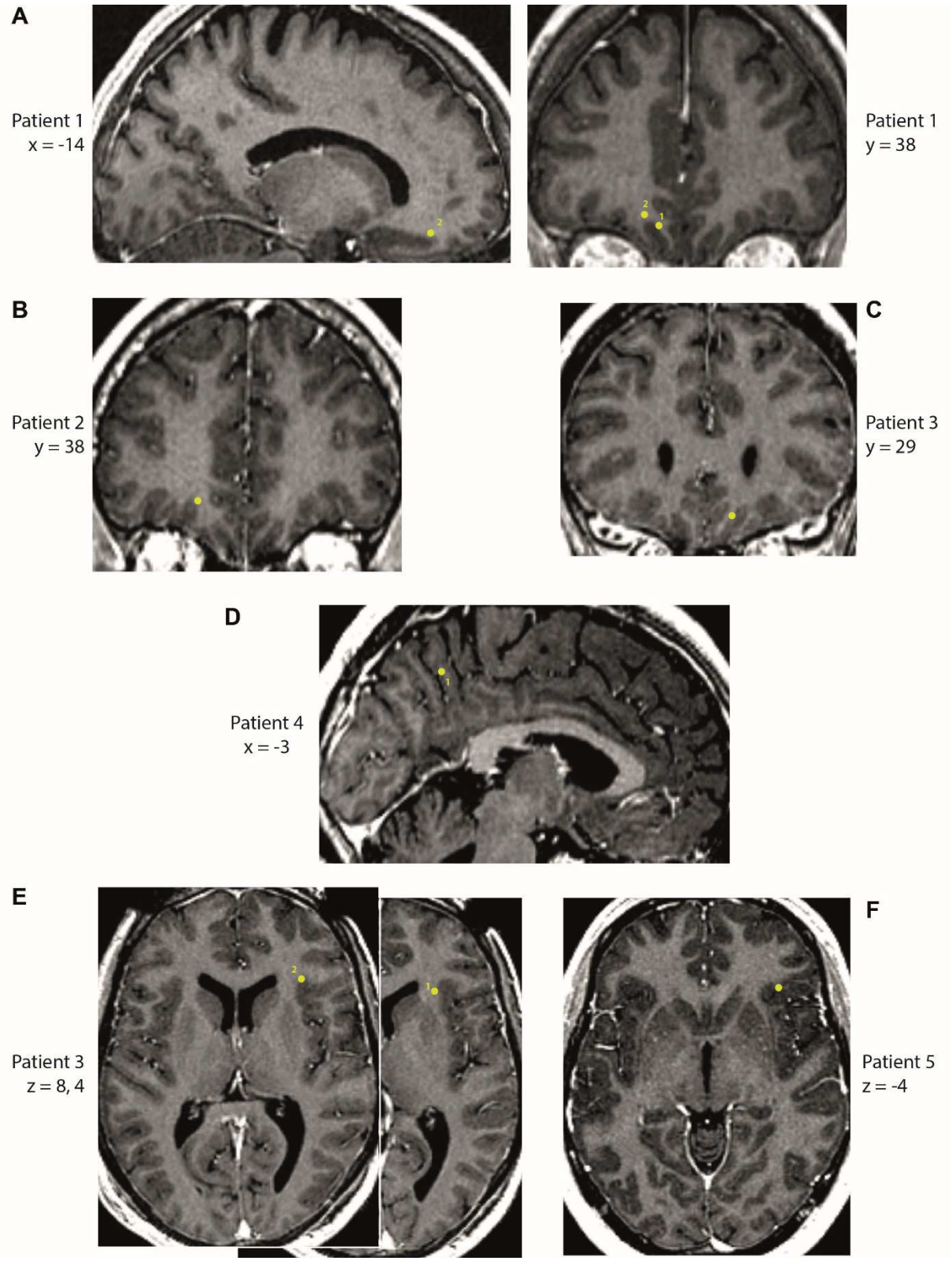
Supplementary material

Control on the cardiac field artefact on MEG data

It is known that MEG and EEG signals are contaminated by the cardiac field artefact [1]. We therefore tested whether the ECG signal differed between “high” and “low” trials on the “I” scale, to check that the effects observed on MEG data were not reflecting a difference in the electrical activity of the heart picked up by the MEG sensors. The ECG signal did not differ (paired *t*-test, “high” vs. “low”, mean ECG amplitude averaged between 387 and 428 ms after the R-peak at each horizontal derivation, all $|t(15)| < 0.81$, all $p > 0.43$; at each vertical derivation, all $|t(15)| < 0.95$, all $p > 0.36$).

This test could not be performed in patients. Although R-peaks were clearly visible in the ECG, the overall level of noise was quite high. However, the cardiac field artefact should be strongly attenuated by the computation of bipolar derivation. Besides, the cardiac-field artefact is not present in the time-window analyzed (300-600 ms post R-peak).

Supplementary Figures



Supplementary Figure 1: Individual MRIs of the patients, normalized to MNI space

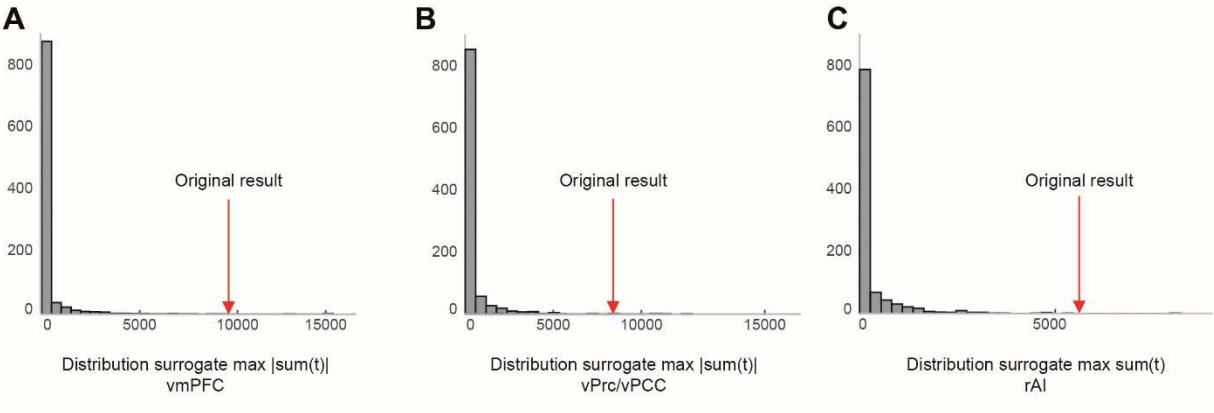
A, Sagittal (left, x=-14) and coronal (right, y=38) views of the MRI of patient 1, showing the recording sites of the electrode shaft located in the vmPFC region.

B, C, Coronal views (y=38, y=29) of the MRIs of patients 2 and 3, respectively, showing the two recording sites in the vmPFC.

D, Sagittal view (x=-3) of the MRI of patient 4, showing the most medial recording site analyzed (recording site 1).

E, Axial views (top: z=8, bottom: z=4) of the MRI of patient 3, showing the two recording sites located in the insular cortex

F, Axial view (z=-4) of the MRI of patient 5, showing the two recording sites in the insular cortex.



Supplementary Figure 2: Test on 1000 surrogate heartbeats

Histograms of the distribution of the maximal cluster *t* statistic obtained for the 1000 permutations of surrogate heartbeats, for iEEG data obtained in recording site 2 of patient 1 (vmPFC, A), recording site 1 of patient 4 (vPrc/vPCC, B) and recording site 2 of patient 5 (rAI, C). The original cluster *t* statistics (red arrows) lie in the tail of the distributions, indicating that the reported effects are truly locked to heartbeats.

Supplementary tables

Supplementary table 1: Information about intracranially recorded patients.

Patient	1	2	3	4	5
Age	39	23	25	30	21
Gender	F	M	F	M	F
Educational level / IQ	Bachelor degree IQ = 106	2 years of higher education IQ = 91	Master degree	2 years of higher education	Bachelor degree
Epilepsy duration (years)	16	20	12	12	13
Epilepsy focus	Left inferior temporal gyrus	Left middle/inferior temporal pole	Right middle temporal gyrus / right hippocampus	Left supplementary motor area	Right anterior/mid cingulate area
Recording system	Neuralynx	Micromed	Neuralynx	Micromed	Neuralynx
Number of clean trials / Total number of trials	27 / 27	33 / 36	26 / 27	45 / 45	24 / 27
Scales analyzed	"Me" Mean: 115±11 Median: 133	"Me" Mean: 101±6 Median: 102	vmPFC: "Me" Mean: 83±12 Median: 72 Insula: "I" Mean: 93±15 Median: 66	"I" Mean: 96±12 Median: 100	"I" Mean: 109±18 Median: 121
Pearson correlation coefficient r, between "I" and "Me" ratings	0.91	0.70	0.84	0.13	0.70
Interbeat interval (ms)	881	771	978	1017	681

Supplementary table 2: anatomical description of the regions present in the reverse inference meta-analysis on the term “self”, based on Automated Anatomical Labeling (AAL) atlas [2]. Only areas with more than 3% of their volume involved are listed.

AAL region	% activation	mm^3	peak z	z/mm^3	MNI		
					X	Y	Z
Left medial orbitofrontal gyrus	28.40	204	7.70	4.86	-8	46	-10
Left posterior cingulate gyrus	24.80	115	6.17	4.60	-4	-52	32
Left superior frontal gyrus, medial	22.30	666	8.73	5.06	-4	54	4
Left anterior cingulate and paracingulate gyri	19.50	273	7.28	4.79	-10	50	2
Right posterior cingulate gyrus	19.10	64	6.18	4.64	2	-52	30
Right medial orbitofrontal gyrus	15.40	132	5.31	4.34	6	48	-8
Right angular gyrus	10.00	176	6.81	4.63	50	-58	28
Left precuneus	9.58	338	7.22	4.58	-2	-58	18
Left angular gyrus	8.95	105	5.97	4.21	-48	-66	32
Right superior frontal gyrus, medial	8.06	172	7.46	4.58	2	54	22
Right gyrus rectus	6.98	52	5.19	4.25	4	44	-16
Left superior frontal gyrus, dorsolateral	6.53	235	6.56	4.42	-8	70	16
Left temporal pole, middle temporal gyrus	6.09	46	5.87	4.51	-52	14	-34
Right precuneus	3.52	115	5.39	4.30	4	-52	24

Supplementary table 3: anatomical description of the regions present in the forward inference meta-analysis on the term "self", based on Automated Anatomical Labeling (AAL) atlas [2]. Only areas with more than 3% of their volume involved are listed.

AAL region	% activation	mm^3	peak z	z/mm^3	MNI		
					X	Y	Z
Left amygdala	88.60	195	16.10	8.13	-22	-4	-16
Right amygdala	63.70	158	15.10	7.85	22	-2	-16
Left lenticular nucleus, pallidum	59.70	175	11.20	5.37	-14	8	-4
Left anterior cingulate and paracingulate gyri	59.10	827	12.30	5.39	-6	46	-4
Left angular gyrus	55.80	655	14.20	5.63	-48	-66	32
Left insula	53.40	992	23.80	7.73	-34	20	0
Right lenticular nucleus, pallidum	51.40	144	9.34	5.23	14	6	0
Left posterior cingulate gyrus	51.40	238	13.40	7.05	0	-54	28
Left medial orbitofrontal gyrus	51.30	369	13.20	6.63	-4	46	-8
Right inferior parietal, excluding supramarginal and angular gyri	49.10	661	10.70	4.80	50	-38	48
Right inferior frontal gyrus, opercular part	44.10	617	11.80	5.31	50	10	28
Left superior frontal gyrus, medial	44.10	1318	15.90	5.98	2	24	44
Right insula	43.20	765	19.70	7.71	34	22	-4
Left inferior frontal gyrus, opercular part	41.30	429	12.60	5.20	-44	6	28
Left thalamus	40.90	450	11.00	5.29	-10	-16	8
Left inferior parietal, excluding supramarginal and angular gyri	40.60	993	8.52	4.50	-34	-56	44
Right caudate nucleus	38.90	387	10.40	5.41	10	10	0
Left inferior frontal gyrus, triangular part	38.30	969	13.70	4.65	-38	22	-2
Left supplementary motor area	37.50	805	17.80	7.33	-2	14	50
Right anterior cingulate and paracingulate gyri	36.70	482	9.34	4.67	4	32	28
Left hippocampus	36.60	341	13.20	5.56	-20	-10	-16
Right medial orbitofrontal gyrus	34.90	299	9.89	5.05	2	48	-8
Left lenticular nucleus, putamen	34.70	350	11.50	4.88	-14	10	-4
Right angular gyrus	34.20	600	11.00	4.73	50	-58	26
Left inferior frontal gyrus, orbital part	33.10	560	13.20	5.62	-34	22	-12
Right posterior cingulate gyrus	32.20	108	14.00	5.85	2	-50	28
Left precentral gyrus	31.80	1121	12.60	4.99	-46	6	32
Right inferior frontal gyrus, triangular part	30.30	651	10.70	4.55	50	22	4
Right hippocampus	27.70	262	11.20	5.35	22	-2	-20
Right thalamus	27.20	287	7.44	4.37	10	-18	4
Right lenticular nucleus, putamen	27.10	288	9.34	4.39	30	18	2
Right median cingulate and paracingulate gyri	25.20	556	13.70	5.59	4	18	44

Left median cingulate and paracingulate gyri	23.20	450	11.80	5.20	-2	14	42
Left supramarginal gyrus	23.20	291	6.35	3.95	-54	-26	20
Right superior frontal gyrus, medial	23.10	493	15.10	5.47	4	24	44
Left precuneus	22.30	787	13.40	5.96	-2	-56	20
Left caudate nucleus	21.40	206	9.89	4.62	-10	10	-2
Right supplementary motor area	21.30	506	13.20	5.97	6	22	46
Right supramarginal gyrus	19.30	380	8.80	4.07	48	-38	44
Left middle temporal gyrus	19.10	943	9.34	4.46	-52	-58	20
Left rolandic operculum	17.80	176	8.52	4.21	-48	6	4
Right inferior frontal gyrus, orbital part	16.80	286	15.90	5.96	36	22	-8
Left superior parietal gyrus	14.80	305	6.89	3.98	-30	-60	44
Left inferior occipital gyrus	14.00	132	6.35	3.92	-42	-74	-4
Right superior parietal gyrus	13.50	301	7.71	4.34	26	-62	56
Right parahippocampal gyrus	13.40	152	8.52	4.12	20	-2	-20
Left parahippocampal gyrus	13.00	127	7.44	3.83	-14	-4	-18
Right precuneus	12.20	398	11.80	4.82	2	-54	28
Right precentral gyrus	11.90	402	14.00	4.93	50	8	32
Right olfactory cortex	11.80	34	6.07	3.93	4	12	-4
Right middle temporal gyrus	11.60	510	7.71	4.02	58	-8	-18
Left postcentral gyrus	11.30	440	8.52	4.28	-38	-22	54
Left middle occipital gyrus	10.70	350	11.00	4.30	-46	-70	4
Left superior frontal gyrus, dorsolateral	8.28	298	7.71	4.27	-24	-4	54
Right gyrus rectus	8.19	61	8.52	4.84	2	50	-16
Right inferior occipital gyrus	8.09	80	5.53	3.86	38	-86	-4
Left gyrus rectus	7.98	68	8.25	4.82	-2	46	-16
Left middle frontal gyrus	7.77	378	8.52	4.04	-26	-2	56
Right middle frontal gyrus	7.46	381	6.62	3.74	28	-2	52
Left calcarine fissure and surrounding cortex	7.44	168	6.89	4.16	-14	-50	4
Right superior temporal gyrus	7.29	229	7.71	3.86	54	-58	22
Right middle occipital gyrus	6.34	133	6.89	3.63	46	-74	4
Right rolandic operculum	5.56	74	7.16	4.03	52	10	0
Right superior occipital gyrus	4.39	62	5.53	3.87	32	-64	40
Left superior temporal gyrus	4.36	100	5.53	3.63	-50	-20	12
Left fusiform gyrus	4.20	97	6.35	3.75	-28	-40	-16
Left temporal pole, superior temporal gyrus	3.66	47	7.16	4.31	-50	6	0
Right hemispheric lobule VI (cerebelum)	3.57	64	4.44	3.48	28	-58	-24
Right fusiform gyrus	3.38	85	4.71	3.49	28	-30	-18
Left olfactory cortex	3.21	9	4.17	3.59	-2	22	-6

Supplementary table 4: iEEG recording sites of interest in the vmPFC region. The recording site showing a significant effect is indicated by a star. The distance between a recording site and a region of interest corresponds to the minimal distance between the recording site and all the voxels belonging to the region.

Patient	**1**		**2**	**3**
Recording site number	1	2*	1	1
MNI coordinates	-9 38 -20	-14 38 -16	-18 38 -12	10 29 -19
Distance to MEG cluster (mm)	9	9	10	11
Distance to region 14m (mm)	4	5	7	Inside right 14m
Distance to region 32 (mm)	13	8	6	/

Supplementary table 5: iEEG recording sites of interest in the vPrc region. The recording site showing a significant effect is indicated by a star. The distance between a recording site and a region of interest corresponds to the minimal distance between the recording site and all the voxels belonging to the region.

Patient	**4**	
Recording site number	1*	2
MNI coordinates	-3 -53 49	-7 -55 52
Distance to MEG cluster (mm)	15	17
Distance to vPrc (mm)	Inside	2

Supplementary table 6: iEEG recording sites of interest in the anterior insula region. The recording site showing a significant effect is indicated by a star. The distance between a recording site and a region of interest corresponds to the minimal distance between the recording site and all the voxels belonging to the region.

Patient	**3**		**5**
Recording site number	1	2*	1
MNI coordinates	28 25 4	32 27 8	36 28 -4
Distance to dAI (mm)	4	4	5
Distance to vAI (mm)	10	14	7

Supplementary Table 7: additional information about MEG recordings, from 16 healthy participants.

	"I" scale	**"Me" scale**
Median ± SEM	135.4±12.4	84.7±11.5
Average number of clean trials ± SEM (total number of trials: 80)	High:40.4±0.2 Low:38.9±0.3	High:40.1±0.2 Low:39.1±0.3
Mean Pearson correlation coefficient between "I" and "Me" ratings (±s.e.m.)	r=0.67±0.04	

Supplementary References

1. Kern, M., Aertsen, A., Schulze-Bonhage, A. & Ball, T. 2013 Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. *Neuroimage* **81**, 178–90. (doi:10.1016/j.neuroimage.2013.05.042)
2. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B. & Joliot, M. 2002 Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289. (doi:10.1006/nimg.2001.0978)

VII. Article III: Imagining the self is associated with neural responses to heartbeats in medial motor regions and the ventromedial prefrontal cortex

A. Abstract in French

Des expériences récentes ont montré le lien entre le soi et le couplage cœur-cerveau, soutenant ainsi les théories selon lesquelles le soi est ancré dans le suivi des signaux internes par le cerveau. Ces expériences ont utilisé des paradigmes fondés sur les pensées spontanées ou sur les illusions corporelles. Ici, nous avons voulu voir si s'imaginer soi-même était aussi associé à ce type de couplage cœur-cerveau. Nous avons demandé à 23 participants de s'imaginer eux-mêmes (depuis une perspective de première personne) ou d'imaginer un ami (depuis une perspective de troisième personne), dans différents contextes. Après avoir imaginé chaque scène, les sujets devaient évaluer à quel point ils avaient réussi à adopter la perspective demandée, à quel point la scène les avait éveillés et la valence de la scène. Nous avons regardé les réponses aux battements cardiaques, mesurées en magnétoencéphalographie, pendant les phases d'imagination. L'amplitude des réponses évoquées aux battements cardiaques n'était pas la même, selon que le sujet s'imaginait lui-même ou imaginait un(e) ami(e). Ces différences s'observaient dans le précuneus antérieur, le cortex cingulaire médian et l'aire motrice supplémentaire (analyse cerveau entier), ainsi que dans le cortex préfrontal ventromédian (analyse par région d'intérêt). L'amplitude des réponses évoquées aux battements cardiaques pendant l'imagination de l'ami(e) corrélait avec la qualité de la perspective. L'amplitude de ces réponses pour le soi corrélait avec la tendance des sujets à se perdre dans leurs pensées dans leur vie quotidienne. Ces résultats montrent que les régions motrices médiales et le cortex préfrontal ventromédian génèrent des réponses évoquées aux battements cardiaques différentielles selon qu'il s'agit de s'imaginer soi-même ou quelqu'un d'autre. Ceci pourrait constituer un mécanisme pour implémenter la distinction soi/autre pendant l'imagination.

B. Article

Imagining the self is associated with neural responses to heartbeats in medial motor regions and the ventromedial prefrontal cortex

Babo-Rebelo Mariana[1], Tallon-Baudry Catherine[1]

[1] Laboratoire de Neurosciences Cognitives (ENS-INSERM), Département d'Etudes Cognitives, Ecole Normale Supérieure-PSL Research University, 75005 Paris, France.

Abstract:

Recent experiments relating the self and heart-brain coupling support theories grounding the self in the neural monitoring of visceral signals. These experiments involved spontaneous thoughts or bodily illusions. We here wanted to see if imagining the self was also associated with such heart-brain mechanism. 23 participants were presented with descriptions of scenarios and had to imagine either themselves from the first-person perspective, or a friend from a third-person perspective. They then evaluated the imagined scene regarding the success in adopting the perspective, as well as valence and arousal. We looked at brain activity, obtained with magnetoencephalography, locked to heartbeats, and compared the amplitude of the obtained Heartbeat-Evoked Responses (HERs) during Self- and Other-imagination. We observed differential HERs between Self and Other, in the anterior precuneus/mid-cingulate/supplementary motor area (whole-brain analysis) as well as in the ventromedial prefrontal cortex (vmPFC, region-of-interest analysis). The amplitude of HERs for Other-imagination correlated with ratings on the success of the perspective-taking. The amplitude for Self-imagination was modulated by the propensity of participants to daydream, as measured by the daydreaming frequency scale. These results show that medial motor regions and the vmPFC generate differential HERs for Self- and Other-imagination. This could constitute a mechanism implementing a Self/Other distinction during imagination.

Introduction

The self has been hypothesized to be anchored in the neural monitoring of visceral signals (Damasio 1999, Park & Tallon-Baudry 2014). This hypothesis received recent experimental support, based on the measure of neural responses to heartbeats. Self-related Heartbeat-Evoked Responses (HERs) were found to co-vary with self-engagement, in tasks exploring the self-relatedness of spontaneous thoughts (Babo-Rebelo et al. 2016a,b) and in tasks targeting the perception of one's own body (Park et al. 2016, Sel et al. 2016). Here we investigate the contrast between self and other, using an imagination task, to test whether HERs are used by the brain to tag the imagined scene as referring to the self or to someone else.

We conducted a mental imagery task, where participants had to imagine either themselves (from a first-person perspective) or a friend (from a third-person perspective) in particular scenarios (Figure 1A) (Ruby & Decety 2001). At each trial, participants were prompted to imagine a scenario from their own first-person perspective, or to imagine a friend. To avoid an imbalance in terms of autobiographical memory between Self and Other, we created a list of scenarios that participants were unlikely to have lived before. These included unreal scenarios (e.g. "At Harry Potter's school"), scenarios in distant or extreme environments (e.g. "In the desert"), or activities participants were unlikely to have performed (e.g. "To participate in a TV show"). After imagining each scenario, subjects were asked to evaluate how well they succeeded in adopting the perspective. They were also asked to report the valence and arousal of the scenarios they imagined, so that we could also assess if the emotions or arousal levels elicited during imagination could modulate HERs.

We tested three main hypotheses: 1) whether HERs distinguish between Self- and Other-imagination; 2) whether the effects take place in the three regions where self-related HERs have been observed so far (ventromedial prefrontal cortex, posterior cingulate/ventral precuneus, right anterior insula or medial motor regions) (Babo-Rebelo et al. 2016a, Park et al. 2016); and 3) whether this effect could be modulated by the vividness of the perspective adopted during the imagined scenarios.

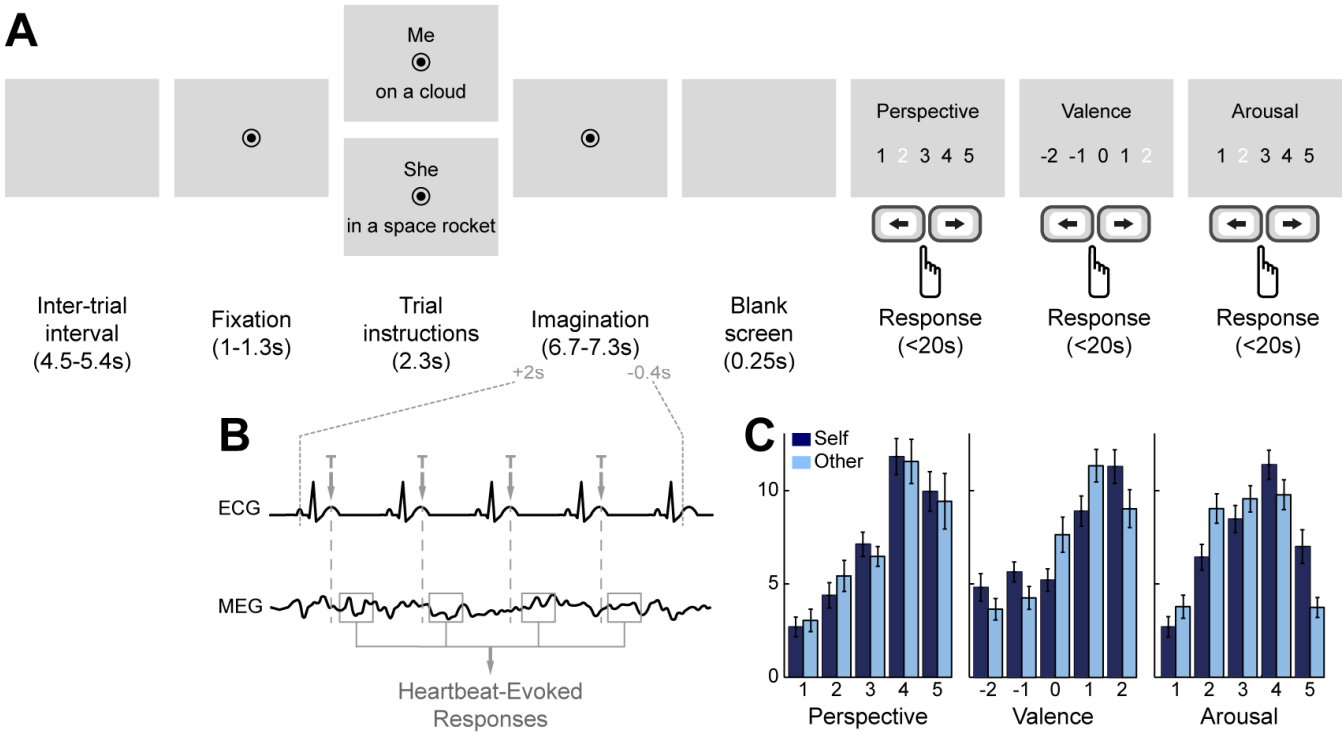


Figure 1: Experimental paradigm and behavior. ***A***, Time course of a trial. Participants had to imagine the person (Self, from the first-person perspective, or Other, from the third-person perspective) in the scenario indicated at each trial, until the fixation disappeared from the screen. They then had to rate the imagined scenario in terms of Perspective (how well they succeeded in adopting the indicated perspective), Valence and Arousal. ***B***, Computation of Heartbeat-Evoked Responses during the imagination period. T-peaks occurring from 2s after the beginning of the imagination period to 0.4s before the end of this period were selected. MEG data was extracted locked to these T-peaks to compute Heartbeat-Evoked Responses. ***C***, Distribution of responses for the Perspective, Valence and Arousal scales, for both Self and Other trials, across all participants. Self trials were significantly more arousing than Other trials (paired t-test on the average Arousal ratings for Self and Other: p=0.0005). Error bars indicate SEM.

Materials & Methods

Participants

25 right-handed volunteers participated in this study after giving written informed consent and were paid for their participation. The study was approved by the local ethics committee. Two participants were excluded from analysis, one because the T-waves were not clear in the electrocardiogram, and the other because of an extremely fast heart-rate (mean interbeat interval = 555ms). Twenty-three participants were thus included in the analysis (9 male; mean age: 24.3±0.6). Participants were screened to exclude cases of prosopagnosia or any cardiac problems. All participants were native French speakers.

Imagination task

Each trial (Fig. 1A) began with a fixation mark (central black dot, radius 0.21° of visual angle, surrounded by a black circle, radius 0.52° of visual angle, on a gray background; 1 to 1.3s), followed by the instruction screen. The person to imagine (condition Self: “Me”, or condition Other: “He”/”She”) was presented above fixation and below the scenario to be imagined. After 2.3s, the instructions were replaced by fixation and subjects had to imagine the scenario. During this period (lasting from 6.7 to 7.3s), participants were instructed to adopt a first-person perspective in trials where they had to imagine themselves, meaning they should imagine the scenario from inside their own body. In trials where they had to imagine the friend, participants had to visualize him/her and not interact with them in the scenario. Participants were also instructed to imagine the person, and not necessarily all the visual details of the scenario. The imagination period stopped with a blank screen (0.25s) and was followed by the presentation of the three scales (the order was randomized between participants). Participants had to rate on 5-point scales: the perspective (how well did you manage to imagine the scenario and to adopt the perspective, from 1: not very well, to 5: very well); the valence (how pleasant was the scenario, from -2: very unpleasant, to 2: very pleasant); and the arousal (how arousing was the imagined scenario, from 1: not arousing, to 5: very arousing) of the imagined scenario. Participants responded by pressing left and right buttons (index and middle finger respectively) to select the appropriate

response. They validated their response with their right thumb, within 20s per scale. A new trial started after an inter-trial interval (blank screen, 4.5 to 5.4s).

Experimental procedure

The day before the experiment, participants were asked to choose the friend they would imagine in the task. The friend had to be the same gender and around the same age as the participant. The participant had to know him/her quite well and had to be able to clearly visualize him/her. It could not be someone the participant was romantically involved with, their best friend or family. To assess the closeness of the selected friend between participants, participants filled in a modified version of the Relationship Closeness Inventory (Berscheid et al. 1989) (RCI, excluding questions related to romantic relationships), where total scores range from 3 to 30. The average RCI score among participants was 12.4±0.8, which is intermediate between close (scores usually around 16) and not close relationships (scores usually around 9).

Before the MEG recording, participants were given written and oral instructions. They performed a short practice block (2 trials of each condition), followed by four blocks of 9 trials of each condition (randomly presented), during which MEG and physiological data were acquired. In the same session, participants also performed a trait-judgment task, a resting state recording and a heartbeat-counting task. These tasks are not reported here. After the recording session, participants completed a short feedback questionnaire. They were then presented with all the scenarios they imagined during the task and had to indicate which scenarios (if any) they had found impossible to imagine, which ones made them laugh, which ones made them anxious, and if they had already lived similar scenarios. Finally, they completed the Daydreaming Frequency Scale (Giambra 1993, Stawarczyk et al. 2012), the Rumination-Reflection Questionnaire (Trapnell & Campbell 1999) and the Trait Anxiety Inventory (Spielberger et al. 1983).

Scenarios

A list of 72 scenarios was created, so that each scenario was presented only once during the experiment. This list was composed of 36 scenarios explicitly related to actions

(examples: to drive a Formula 1 car, to build a standing stone), and 36 scenarios not directly related to actions (examples: in the Middle Ages, in the jungle). Some scenarios involved other people and possibly social interactions (examples: to shake hands with Obama, at the Rio carnival).

Scenarios were randomly assigned to each condition, with action- and non-action-related scenarios equally distributed between conditions. Scenarios assigned to condition Self for subject 1 were assigned to condition Other for subject 2 and vice-versa, for all pairs of subjects. This way, each scenario was associated with "Self" and with Other" conditions the same number of times across subjects. There was no significant difference in the number of characters contained in the scenario descriptions between conditions Self and Other ($t_{(22)}$=-0.1, p=0.9).

Recordings

Continuous magnetoencephalographic (MEG) data were acquired using a whole-head MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX, sampling rate of 1000Hz, online low-pass filtered at 330Hz). Electrocardiogram data (EKG, 0.03-330Hz) were obtained from 7 electrodes placed around the base of the neck and referenced to a left abdominal location. The ground electrode was located on the back of the neck. Two EKG electrodes were placed over the left and right clavicles, two over the top of the left and right shoulders, two over the left and right supraspinatus muscle and one over the upper part of the sternum. The electrodermal activity (EDA, two electrodes on the sole of the left foot) as well as respiratory activity (respiratory belt positioned around the chest, at the level of armpits, respiratory transducer TSD201 BIOPAC system) were also recorded (low-pass filtered at 330Hz). Electromyographic activity (EMG, two electrodes on the right cheek, 10-330Hz) from the right zygomaticus major was acquired in order to control for facial muscle activity (laughter, in particular). Horizontal and vertical eye position and pupil diameter were monitored using an eye-tracker (EyeLink 1000, SR research) and recorded together with MEG, EKG, EDA and respiratory data. Stimuli were presented on a semi-translucent screen at 85 cm viewing distance.

MEG data preprocessing

Continuous MEG data were denoised using temporal signal space separation (TSSS, as implemented in MaxFilter) and filtered between 0.5 and 40Hz (4^{th} order Butterworth filter). Blinks and saccades larger than 2 degrees were identified by the Eyelink system. Epochs contaminated by large movement or muscle artefacts were visually detected. Independent Component Analysis (ICA) as implemented in the Fieldtrip toolbox (Oostenveld et al. 2011) was used to correct for the cardiac field artifact, on both magnetometers and gradiometers, based on epochs of -0.2 to 0.2s around the R-peaks of interest devoid of movement, muscle, blink or saccade artefacts. Because TSSS induces rank-deficiency, we defined the number of ICA components by first computing a Principal Component Analysis. We then removed all independent components with mean pairwise phase consistency (Vinck et al. 2010) with the ECG in the 0-25 Hz range larger than two standard deviations of all components. We iterated this procedure until no outlier components were found or a maximum of two excluded components was reached. ICA corrected MEG data were then low-pass filtered at 30 Hz (4^{th} order Butterworth filter).

Heartbeat-evoked responses (HERs)

We considered that ratings for which the reaction time was larger than 2.5 standard deviations of the reaction time of the corresponding scale revealed unsuccessful imagination. The corresponding trials were excluded from analysis.

We first detected the R-peaks, by correlating the ECG with a template QRS complex defined on a subject-by-subject basis and identifying the local maximum within the episodes of correlation larger than 0.7. T-peaks were then detected by first, correlating the ECG with a template of the T-peak; second, identifying the local maxima within episodes of correlations above a certain correlation value (adapted for each subject) that followed an R-peak by at most 0.4s. R- and T-peak detection was visually verified in all subjects. The T-peaks occurring during the imagination period (from 2 seconds after the beginning of the period, to -0.4s before the end of the period) were used for HER computation. By excluding the beginning of the imagination period, we make sure subjects already started imagining the scenario. We rejected epochs (from 0.1s before to 0.4s after the selected T-peaks) contaminated with

saccades larger than 2° of visual angle from fixation, blinks and movement or muscular (in particular of the zygomaticus) artifacts. Artefact-free HERs corresponding to Self and Other trials were computed by averaging magnetometer data across heartbeats, from 0.1s before the T-peak to 0.4s after the T-peak.

Statistical analyses

The difference in HERs between Self and Other was tested on magnetometers, in the artefact-free time window 80-350ms, after the T-peak, using a cluster-based permutation t-test (Maris & Oostenveld 2007). This method does not require any a priori on spatial regions or latencies thereby correcting for multiple comparisons in time and space. A paired *t*-test was performed to compare HERs for Self versus Other. Individual samples whose *t*-value was below a threshold ($p<0.05$, two-tailed) were clustered together based on temporal and spatial adjacency (with a minimum of 4 neighboring channels/time). A cluster was characterized by the sum of the *t*-values of the individual samples. To test whether such a cluster could be obtained by chance, we permuted the labels "Self" and "Other" 10,000 times and selected the maximal positive cluster-level statistic and the minimal negative cluster-level statistic at each randomization. The two-tailed Monte-Carlo *p*-value corresponds to the proportion of elements in the distribution of maximal (or minimal) cluster-level statistics that exceeds (or is inferior to) the originally observed cluster-level test statistics. The amplitude of the cluster corresponds to the average of magnetometer data across the sensors and time window showing a significant difference.

Bayes Factors were computed to evaluate evidence in favor of the null hypothesis, in both paired t-tests and two-sample t-tests, with the Jeffreys-Zellner-Siow prior, as implemented in the online calculator tool (http://pcl.missouri.edu/bayesfactor) (Kass & Raftery 1995, Liang et al. 2008, Wetzels & Wagenmakers 2012).

Surrogate heartbeats

To test whether the observed effects were only locked to heartbeats, we checked whether differences between Self and Other trials could be obtained with a sampling of neural data that was unsynchronized with heartbeats. We created 1,000 permutations of

heartbeats, where the timing of the heartbeats of trial *i* in the original data was randomly assigned to trial *j*. The same criteria for rejecting artefactual epochs and computing HERs were applied. For each permutation, we obtained a set of neural responses to surrogate heartbeats and computed the cluster summed *t*-statistics as described above. For each permutation we extracted the smallest negative sum of *t*-values, and compared the distribution of those surrogate values with the observed original sum of *t*-values.

Anatomical MR acquisition and preprocessing

An anatomical T1 scan was acquired for 22 participants. Segmentation of the data was processed with automated algorithms provided in the FreeSurfer software package (Fischl et al. 2004) (http://surfer.nmr.mgh.harvard.edu/). Segmentations were visually inspected and edited when necessary. The white-matter boundary was determined using FreeSurfer and was used for subsequent minimum-norm estimation.

Source reconstruction

We reconstructed sources of HERs occurring from 2 to 4s after the onset of the imagination period. Source reconstruction and surface visualization were performed with the BrainStorm toolbox (Tadel et al. 2011). For the participant who did not have an anatomical scan, we warped the ICBM152 anatomical template (http://bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009) to fit the shape defined by the digitized head points obtained before MEG acquisition. After co-registration between the individual anatomy and MEG sensors, cortical currents were estimated using a distributed model consisting of 15,002 current dipoles from the combined time series of magnetometer and gradiometer signals using a linear inverse estimator (weighted minimum-norm current estimate, signal-to-noise ratio of 3, Whitening PCA, depth weighting of 0.5) in an overlapping-spheres head model. Dipole orientations were constrained to the individual MRIs. Cortical currents were then averaged over the time windows for which a significant difference between Self and Other was identified in sensor space, spatially smoothed (FWHM 7mm) and projected to a standard brain model (ICBM152, 15,002 vertices).

Reliable differences in dipole current values were identified using the cluster-based procedure (first-level p-value: 0.05, 2 neighboring sensors) as described for the sensor level analysis applied to the 15,002 vertices. The obtained Monte-Carlo p-value was corrected for multiple comparisons over space.

Region of interest analyses

To compare the current results with previous work (Babo-Rebelo et al. 2016a,b), we analyzed the ventromedial prefrontal region (conjunction of p32, 11m and 14m (Neubert et al. 2015)), the posterior medial cortex (conjunction of ventral precuneus and ventral posterior cingulate cortex (Bzdok et al. 2015)) and the right anterior insula (conjunction of the dorsal and ventral parts (Deen et al. 2011)), using masks provided by the respective authors. The masks were transformed to fit the anatomical template of Brainstorm, using the function ImCalc of SPM12. The mean time-course in each ROI was extracted using Brainstorm (masks loaded as scouts, and default sign-flipping option on). We then tested each ROI time-course for differences between Self and Other, in a time window 80-350ms after the T-peak, using the clustering procedure described above.

Arousal-related measures

Interbeat intervals consisted of the average time distance between the T-peaks in the imagination period and the heart rate variability corresponded to the standard deviation of the interbeat intervals.

Blinks were automatically detected with the Eyelink software. For pupil diameter computation, the corresponding time windows were extended by 80ms on each side. We also identified and rejected all variations in pupil diameter > 100 (arbitrary units) in a 300ms time window. To analyze pupil diameter, portions of data containing blinks were linearly interpolated and a low-pass fourth-order Butterworth filter at 10Hz was applied. Data were then epoched from 2 to 4 seconds after the onset of the imagination period. Epochs with >30% noisy data were excluded from analysis. The remaining epochs were z-scored. One subject was excluded from pupil diameter analysis for having a very low number of clean trials (n=34, < 3 SDs less than the average number of trials in the other participants).

To compute alpha power, ICA-corrected MEG data were bandpass filtered between 8 and 12Hz (4^{th} order Butterworth filter) and the corresponding alpha-band power was computed using the Hilbert transform. Data from the 15 sensors showing the largest alpha power at the group level were averaged. Portions of data contaminated by blinks were excluded from analysis.

Results

Behavioral results

The distribution of ratings on the Perspective, Valence and Arousal scales is presented in Figure 1C. Mean ratings did not differ between Self and Other conditions in the Perspective or Valence scales (Perspective: mean Self: 3.6±0.1 SEM, mean Other: 3.5±0.1, paired t-test Self x Other, $t_{(22)}$=0.7, p=0.5; Valence: Self: 3.5±0.1, Other: 3.5±0.1, $t_{(22)}$=-0.5, p=0.6; uncorrected p-values). Imagining oneself was rated as being more arousing than imagining the other (Arousal: Self: 3.4±0.1, Other: 3±0.1, $t_{(22)}$=4.1, p=0.0005; uncorrected p-value).

HER amplitude differs between self and other

We compared the amplitude of HERs occurring during imagination of self with the amplitude of HERs occurring during imagination of other, from 2s after the imagination period onset to -0.4 seconds before the end of the imagination period (Fig. 1B). HERs significantly differed over posterior sensors (Fig. 2A), in the time window 313-328ms after the T-peak (Fig. 2B; cluster sum(t)=-555.0, Monte-Carlo p=0.029).

To show that this effect was truly locked to heartbeats and not driven by slow fluctuations of neural activity differing between conditions, we permuted heartbeat timings between trials 1,000 times and performed the same analyses on these surrogate heartbeats. Only 15/1,000 permutations led to a cluster *t* statistic larger than the original one (Monte-Carlo p=0.015, one-sided), which demonstrates that our effect is truly an evoked-response to heartbeats.

We then looked at the temporal evolution of the effect, during the imagination period. The difference in HERs between Self and Other was larger in the time window 2 to 4 seconds from the onset of the imagination period and both Self and Other cluster amplitudes differed from zero (paired t-test between the cluster amplitude for Self and the cluster amplitude for Other: $t_{(22)}$=-3.2, p=0.0037; t-test cluster amplitude against zero: Self: $t_{(22)}$=-2.33, p=0.029, Other: $t_{(22)}$=2.58, p=0.017) (Fig. 2C). The difference was not significant in the time window 4 to 6 seconds (paired t-test: $t_{(22)}$=-1.19, p=0.25; t-test cluster amplitude against zero: Self: $t_{(22)}$=0.35, p=0.73, Other: $t_{(22)}$=1.86, p=0.077). In the following, we concentrate our analyses on T-peaks occurring in the time window 2 to 4 seconds from the onset of the imagination, where effect size is maximal.

We tested whether HERs co-varied with the ratings on the Perspective, Valence and Arousal scales. For each scale and for each subject we correlated the trial-by-trial cluster amplitude with the corresponding rating and compared the correlation coefficients between Self and Other across subjects. Correlation coefficients with Valence did not differ between Self and Other (mean Fisher z-transformed Pearson correlation coefficients, Self: r=0.024±0.029, Other: r=-0.031±0.025; paired t-test: $t_{(22)}$=1.56, p=0.39, Bonferroni corrected for the three scales tested) nor correlation coefficients with Arousal (Self: r=0.004±0.025, Other: r=-0.018±0.025; paired t-test: $t_{(22)}$=0.80, p=1, Bonferroni corrected). For the Perspective scale, correlation coefficients for Self differed from correlation coefficients for Other (Fig. 2D; Self: r=0.027±0.021, Other: r=-0.057±0.024; paired t-test: $t_{(22)}$=2.92, p=0.024, Bonferroni corrected). More precisely, this difference corresponded to a significant negative correlation between cluster amplitude for Other and ratings on the Perspective scale (t-test against zero of the Fisher z-transformed Pearson correlation coefficients: $t_{(22)}$=-2.37, p=0.027). The correlation did not differ from zero for Self ($t_{(22)}$=1.23, p=0.21).

To summarize, our results show that HERs differ between Self and Other, and that during imagination of the friend, their amplitude is parametrically modulated by how well the friend is visualized.

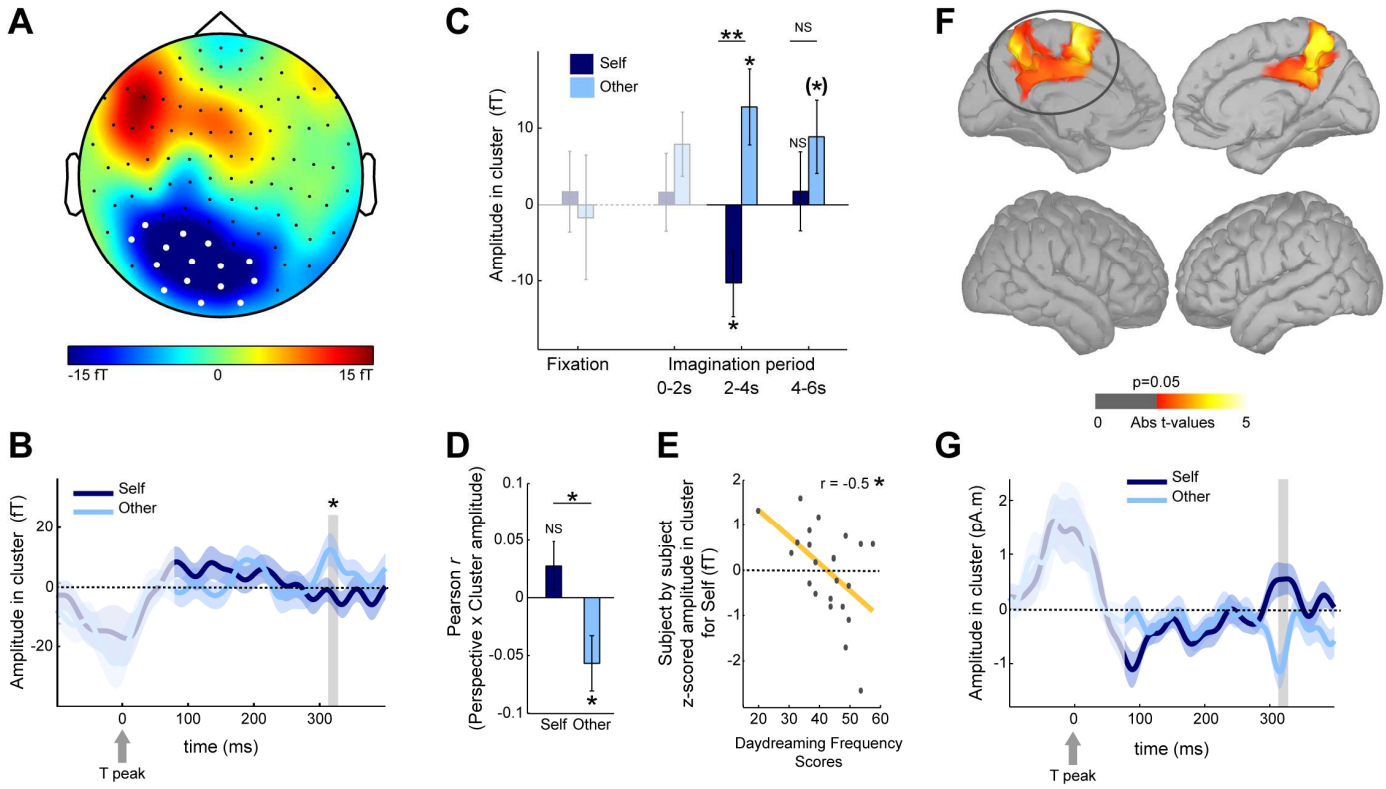


Figure 2: Differential HERs for imagining oneself or someone else. ***A***, Topographical map of the HER difference between "Self" and "Other" conditions, grand-averaged across 23 participants, in the 313-328ms time window in which a significant difference was observed (Mont-Carlo p=0.029). White dots represent the sensors contributing to the significant cluster. ***B***, Time course of the HER (±SEM) for "Self" and "Other", averaged over the white sensors indicated in A. The signal that might be residually contaminated by the cardiac artifact appears in lighter color. The grey area represents the time window in which a significant difference was observed. ***C***, Temporal evolution of the effect, during the imagination period. Amplitude in cluster corresponds to the average brain activity in the time window and sensors revealing a significant effect. Cluster amplitude was computed during fixation (1-1.3s), and during the imagination period divided in three windows of 2 seconds (0-2s, 2-4s, 4-6s). The largest cluster amplitude differences between Self and Other were observed in the window 2-4s. ***D***, Mean Pearson correlation coefficient *r*, across subjects, between cluster amplitude and Perspective ratings. ***E***, Pearson correlation between the z-scored mean amplitudes in the cluster in the condition Self and the Daydreaming Frequency Scores. Each dot represents one subject. ***F***, Neural sources of the differential HERs found in medial motor regions. The left anterior precuneus, mid-cingulate cortex and supplementary motor area survived correction for multiple comparisons (circled, Monte-Carlo p=0.010). The right homologous region did not reach significance (Monte-Carlo p=0.060; threshold for visualization: >30 contiguous vertices at uncorrected p<0.05). ***E***, Time course of the HERs (±SEM) in the region circled in F. The grey area represents the time window that is significant at the sensor level. NS: non-significant; *: p<0.05; **: p<0.01.

Modulation by personality factors

We tested whether the amplitude of the effects was modulated by personality traits assessed via self-report questionnaires, in particular the propensity to daydream during daily life (Daydreaming Frequency Scale, (Giambra 1993)) and trait anxiety (Spielberger et al. 1983). We correlated the effect size (difference between cluster amplitude for Self and

cluster amplitude for Other, on heartbeats occurring between 2 and 4s from the onset of the imagination period, z-scored) with ratings on these questionnaires.

No correlation was found between the effect size and anxiety scores (Pearson correlation $r_{(21)}$=0.064, p=0.77, FDR corrected for the two scales tested), whereas the effect size was correlated with daydreaming frequency scores (Pearson correlation $r_{(21)}$=-0.51, p=0.024, FDR corrected). More specifically, this result was driven by the cluster amplitude for Self, which significantly correlated with the daydreaming frequency scores (Fig. 2E, $r_{(21)}$=-0.52, p=0.012, uncorrected). Subjects who are prone to daydreaming in daily life have large HER amplitude when imagining themselves. On the contrary, cluster amplitude for Other did not correlate with daydreaming frequency scores ($r_{(21)}$=0.27, p=0.22, uncorrected).

HERs in medial motor regions are responsible for these effects

To identify the regions generating the differential HERs, we reconstructed HER sources for Self and Other, averaged the reconstructed neural currents in the time window where we found an effect (313-328ms after the T-peak) and performed a cluster-based permutation test over all 15,002 vertices to compare activations for Self and Other. The differential HER amplitude was located in medial motor regions, comprising the anterior precuneus, the mid-cingulate cortex and the left supplementary motor area (Fig. 2F, 2G). This difference was significant on the left hemisphere (Table 1, cluster sum(t)=1,125, Monte-Carlo p=0.010), but it did not reach significance and did not include the supplementary motor area on the right (Fig. 2F, cluster sum(t)=-743, Monte-Carlo p=0.060). Additionally, these regions overlapped with regions involved in autonomic regulation (Fig. 3), according to a meta-analysis of Beissner *et al.* (Beissner et al. 2013).

Other regions did not survive the stringent correction for multiple comparisons applied here, namely the bilateral ventromedial prefrontal cortex and the right superior frontal gyrus (Supplementary Figure).

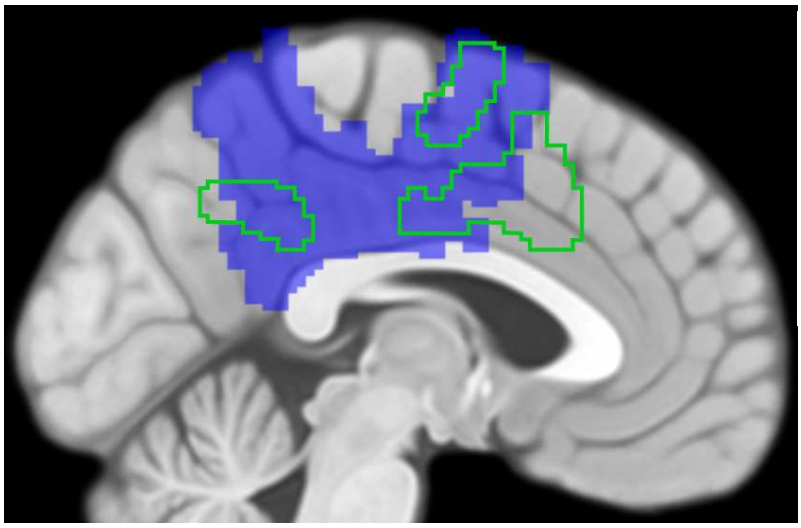


Figure 3: Differential HERs and autonomic regulation. Regions in blue showed differential responses to heartbeats during Self- vs Other-imagination. Green outlines correspond to a meta-analysis on autonomic regulation (Beissner et al. 2013). Sagittal view, x = -3.

AAL regions	sum(*t*)	peak *t*	MNI coordinates (peak *t*)		
			X	Y	Z
Left Supplementary Motor Area	177.6	4.8	-2	-6	60
Left mid-Cingulate	107.7	4.0	-10	-4	40
Left Precuneus	118.5	4.2	-6	-56	50

Table 1: Anatomical description of the main regions showing significant differential HERs (Fig. 2F).

Region of interest analysis

In the spontaneous thought paradigm previously studied, we observed self-related HERs in the posteromedial cortex (PMC) and in the ventromedial prefrontal cortex (vmPFC) in a whole-brain analysis, as well as in the right anterior insula (rAI) in a region of interest (ROI) approach. We thus tested HERs for differences between Self and Other in these three ROIs. In the PMC, a candidate cluster was found in the time window 302-321ms after the T-peak, but the difference between Self and Other was not significant (Fig. 4A, sum(t)=44.9, Monte-Carlo p=0.72, Bonferroni corrected for the three ROIs tested). In the vmPFC, we found a significant difference in HER amplitude, in the time window 260-339ms after the T-peak (Fig. 4B, sum(t)=260.6, Monte-Carlo p=0.012, Bonferroni corrected). The clustering procedure did not return any candidate cluster for the rAI (Fig. 3C, p=1).

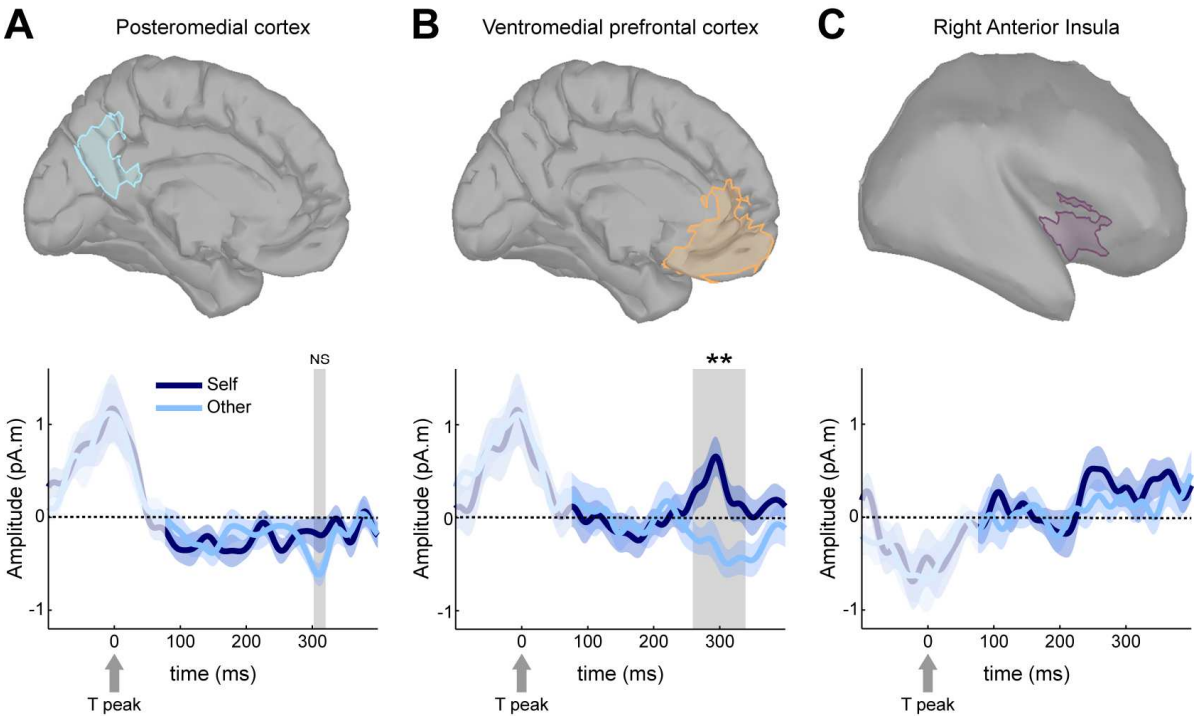


Figure 4: Region of interest analysis of the Posteromedial cortex (***A***), the Ventromedial prefrontal cortex (***B***) and the Right anterior insula (***C***). The areas outlined in the three inflated brains correspond to the vertices included in each of the regions of interest, from which we extract the corresponding average time course. The signal that might be residually contaminated by the cardiac artifact appears in lighter color and was not included in this analysis. The grey areas indicate time windows where a candidate cluster was found by the clustering procedure. NS: non-significant; **: p<0.01.

Control for arousal effects

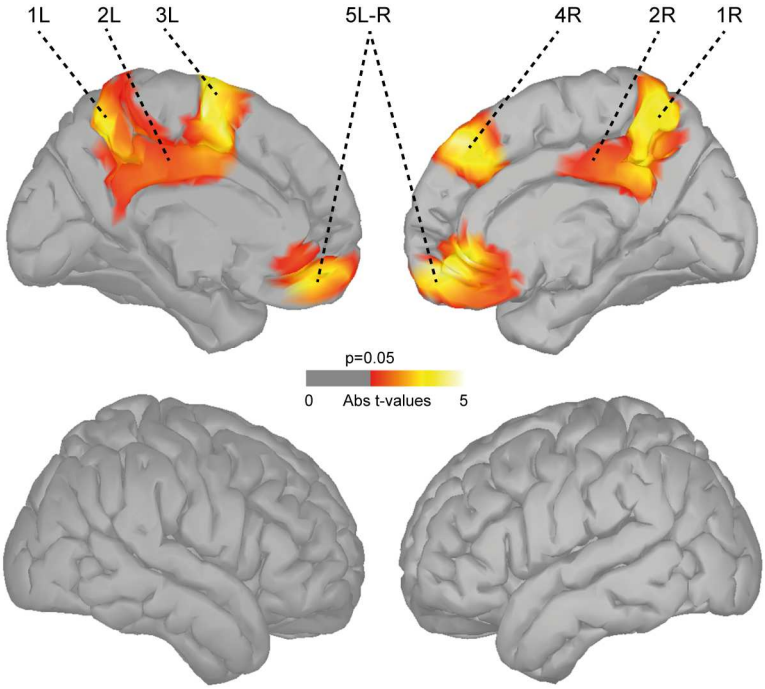
Self trials were judged as being more arousing than Other trials. Although the effect size did not correlate with Arousal ratings, we additionally controlled other parameters related to arousal.

First, we computed the average pupil diameter in the time window 2-4s after the onset of the imagination period. Pupil diameter did not differ between Self and Other (mean Self=4.1±2.4 a.u., Other=-4.1±2.5 a.u.; paired t-test: $t_{(21)}$=1.65, p=0.11; Bayes Factor: 1.40, anecdotal evidence in favor of the null).

Moreover, we observed no difference between Self and Other, in neither the interbeat interval (mean IBI Self = 852.9±24.5ms, mean IBI Other = 852.4±24.2ms; paired t-test: $t_{(22)}$=0.21, p=0.83; Bayes Factor = 4.48, substantial evidence in favor of the null) nor the

heartrate variability (mean HRV Self = 48.9±2.5, mean HRV Other = 52.8±3.7; paired t-test: $t_{(22)}$=-1.67, p=0.11; Bayes Factor = 1.37, anecdotal evidence in favor of the null).

We then looked at the average alpha power in the same time window of the imagination period, as an index of the arousal level. Alpha power did not differ between Self and Other (mean Self = 5.89±1.15 fT^2Hz^{-1}, mean Other = 5.95±1.17 fT^2Hz^{-1}; paired t-test: $t_{(22)}$=-1.49, p=0.15; Bayes Factor = 1.73, anecdotal evidence in favor of the null).



Supplementary Figure:
Regions responding differentially to heartbeats between Self and Other, in the time-window 313-328ms after the T-peak, with a more liberal threshold than in Fig. 2F (uncorrected p<0.05, >75 contiguous vertices). The numbers refer to:
1) anterior precuneus,
2) mid-cingulate cortex,
3) supplementary motor area,
4) superior frontal gyrus,
5) ventromedial prefrontal cortex.
L: left, R: right.

Discussion

We here show that the amplitude of heartbeat-evoked responses (HERs) in medial motor regions (anterior precuneus, mid-cingulate and supplementary motor area - SMA) differed between imagination of Self and imagination of a friend. A region of interest approach further showed differential HERs in the ventromedial prefrontal cortex (vmPFC), consistent with previous findings on the self-relatedness of spontaneous thoughts (Babo-Rebelo et al. 2016a,b). Two other regions of interest, in the posterior cingulate / ventral precuneus (PCC/vPrc) and in the right anterior insula (rAI), did not display differential HERs for the imagination of Self and Other. In addition, the amplitude of HERs for Self correlated with daydreaming scores, suggesting a link between spontaneous cognition in daily life and the more constrained experimental setting proposed here. The amplitude of HERs for Other was negatively correlated with the vividness of the third-person perspective adopted during imagination. Even though imagination of Self was rated as being more arousing, we did not find any significant evidence for a difference in arousal levels between conditions in either pupil diameter, alpha power, interbeat interval or heartrate variability. We here demonstrate that Self- and Other-imagination are associated with differential HERs. These results thus generalize the role of HERs in implementing the self by showing HER involvement in the Self vs Other distinction during mental imagery, beyond spontaneous cognition (Babo-Rebelo et al. 2016a,b) or passive exposure to bodily illusions (Park et al. 2016, Sel et al. 2016).

We hypothesized that HERs distinguishing between Self and Other would take place in regions where a link between HERs and the self has already been shown, namely the vmPFC, the PCC/vPrc, the rAI (Babo-Rebelo et al. 2016a,b) and medial motor regions (Park et al. 2016). The largest effect was here observed in a whole-brain analysis in midline motor regions. The involvement of the SMA and mid-cingulate motor cortex (Dum et al. 2009, Naito et al. 2016), is consistent with the fact that half of the scenarios explicitly referred to motor actions (ex: “to sheer sheep”). Subjects mentally simulated actions, in the Self condition, and visualized their friend’s actions, in the Other condition. Given the known role of the SMA in distinguishing between simulation and observation of actions (Macuga & Frey 2012, Zentgraf et al. 2005), differential HERs in the SMA could thus be interpreted as encoding simulation vs observation of actions. Motor actions were here imagined from a

first- or third-person perspective. This difference in perspective taking has been associated with the anterior precuneus (Ruby & Decety 2001), where we find differential HERs. The remaining scenarios referred to contexts where participants had to navigate (example: “in the desert”). This is compatible with the known role of both the SMA and the anterior precuneus in spatial navigation (Huang & Sereno 2013). Moreover, recent findings showed a link between HERs in the SMA and mid-cingulate cortex and bodily self-consciousness (Park et al. 2016), suggesting that there was here a bodily component in the imagined scenarios. Visceral signals could thus be integrated in these medial motor regions to contribute to a body-centered reference frame (Bernier & Grafton 2010), defining the self, and allowing a distinction between Self and Other in mental imagery.

In the present experiment, a region of interest analysis revealed that HERs also differed between Self and Other in the vmPFC, where HERs have been previously associated with thinking about oneself (Babo-Rebelo et al. 2016a). However, contrarily to our initial hypotheses, we did not find any differential HERs in the PCC/vPrc nor in the rAI, which have been found to characterize non-reflective aspects of the self (Babo-Rebelo et al. 2016a,b). Here, both Self and Other conditions involve a non-reflective self; a subject that is acting – in the Self condition, or a subject that is observing – in the Other condition. Maybe, these non-reflective aspects of the self are better explored in gradients of self-involvement, as previously studied in spontaneous thoughts (Babo-Rebelo et al. 2016a,b), rather than in the Self vs Other contrast analyzed here.

We also hypothesized that HER amplitudes could be modulated by the vividness of the perspective adopted during the imagined scenarios. We found that HER amplitude during imagination of Other was negatively correlated with the vividness of the perspective. The direction of this correlation was unexpected, but coherent with work on the relationship between vividness of imagery and gastrointestinal activity (Vianna et al. 2009) or heart rate variability (Laor et al. 1999). Surprisingly, we found no modulations of HER amplitude for Self. However, our analysis was restricted to sensors and latencies where HERs differed between Self and Other. Further analyses should focus on exploring the correlation between HERs and ratings of perspective, in particular in the posterior cingulate / precuneus / retrosplenial cortex (Cabeza & St Jacques 2007, Dijkstra et al. 2017, Richter et al. 2016).

Acknowledgments: We thank Margaux Romand-Monnier and Christophe Gitton for help with data acquisition, Anne Buot for help with data preprocessing and Stephen Whitmarsh for useful comments on previous versions of the manuscript.

References

Babo-Rebelo M, Richter C, Tallon-Baudry C. 2016a. Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. *J. Neurosci.* 36(30):7829–40

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C. 2016b. Is the cardiac monitoring function related to the self in both the default network and right anterior insula? *Philos. Trans. R. Soc. B*. 371(1708):1–13

Beissner F, Meissner K, Bär K-J, Napadow V. 2013. The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. *J. Neurosci.* 33(25):10503–11

Bernier PM, Grafton ST. 2010. Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. *Neuron*. 68(4):776–88

Berscheid E, Snyder M, Omoto AM. 1989. The relationship closeness inventory : Assessing the closeness of interpersonal relationships. *J. Pers. Soc. Psychol.* 57(5):792–807

Bzdok D, Heeger A, Langner R, Laird AR, Fox PT, et al. 2015. Subspecialization in the human posterior medial cortex. *Neuroimage*. 106:55–71

Cabeza R, St Jacques P. 2007. Functional neuroimaging of autobiographical memory. *Trends Cogn. Sci.* 11(5):219–27

Damasio AR. 1999. *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. New York: Harcourt

Deen B, Pitskel NB, Pelphrey K a. 2011. Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex*. 21(7):1498–1506

Dijkstra N, Bosch S, van Gerven MAJ. 2017. Vividness of visual imagery depends on the neural overlap with perception in visual areas. *J. Neurosci.* 37(5):1367–73

Dum RP, Levinthal DJ, Strick PL. 2009. The spinothalamic system targets motor and sensory areas in the cerebral cortex of monkeys. *J. Neurosci.* 29(45):14223–35

Fischl B, Van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, et al. 2004. Automatically parcellating the human cerebral cortex. *Cereb. Cortex*. 14(1):11–22

Giambra LM. 1993. The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Exp. Gerontol.* 28(4–5):485–92

Huang R-S, Sereno MI. 2013. Bottom-up retinotopic organization supports top-down mental imagery. *Open Neuroimag. J.* 7(858):58–67

Kass RE, Raftery AE. 1995. Bayes Factors. *J. Am. Stat. Assoc.* 90(430):773–95

Laor N, Wolmer L, Wiener Z, Sharon O, Weizman R, et al. 1999. Image vividness as a psychophysiological regulator in posttraumatic stress disorder. *J. Clin. Exp. Neuropsychol.* 21(1):39–48

Liang F, Paulo R, Molina G, Clyde MA, Berger JO. 2008. Mixtures of g-priors for Bayesian variable

selection. *J. Am. Stat. Assoc.* 103(481):410–23

Macuga KL, Frey SH. 2012. Neural representations involved in observed, imagined, and imitated actions are dissociable and hierarchically organized. *Neuroimage*. 59(3):2798–2807

Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods*. 164(1):177–90

Naito E, Morita T, Amemiya K. 2016. Body representations in the human brain revealed by kinesthetic illusions and their essential contributions to motor control and corporeal awareness. *Neurosci. Res.* 104:16–30

Neubert F-X, Mars RB, Sallet J, Rushworth MFS. 2015. Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 1–10

Oostenveld R, Fries P, Maris E, Schoffelen JM. 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011(156869):

Park H-D, Bernasconi F, Bello-Ruiz J, Pfeiffer C, Salomon R, Blanke O. 2016. Transient modulations of neural responses to heartbeats covary with bodily self-consciousness. *J. Neurosci.* 36(32):8453–60

Park H-D, Tallon-Baudry C. 2014. The neural subjective frame: from bodily signals to perceptual consciousness. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369(1641):20130208

Richter FR, Cooper RA, Bays PM, Simons JS. 2016. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *Elife*. 5:e18260

Ruby P, Decety J. 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4(5):546–50

Sel A, Azevedo RT, Tsakiris M. 2016. Heartfelt self: cardio-visual integration affects self-face recognition and interoceptive cortical processing. *Cereb. Cortex*. 1–12

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. 1983. *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press

Stawarczyk D, Majerus S, Van der Linden M, D'Argembeau A. 2012. Using the Daydreaming Frequency Scale to Investigate the Relationships between Mind-Wandering, Psychological Well-Being, and Present-Moment Awareness. *Front. Psychol.* 3(September):363

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM. 2011. Brainstorm: A user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011:879716

Trapnell PD, Campbell JD. 1999. Private Self-Consciousness and the Five-Factor Model of Personality: Distinguishing Rumination From Reflection. *J. Pers. Soc. Psychol.* 76(2):284–304

Vianna EPM, Naqvi N, Bechara A, Tranel D. 2009. Does vivid emotional imagery depend on body signals? *Int. J. Psychophysiol.* 72(1):46–50

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CMA. 2010. The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. *Neuroimage*. 51(1):112–22

Wetzels R, Wagenmakers E-J. 2012. A default Bayesian hypothesis test for correlations and partial correlations. *Psychon. Bull. Rev.* 19:1057–64

Zentgraf K, Stark R, Reiser M, Künzell S, Schienle A, et al. 2005. Differential activation of pre-SMA and SMA proper during action observation: Effects of instructions. *Neuroimage*. 26(3):662–72

VIII. General discussion

A. Main results and discussion on the consistency between tasks

1. HERs encode the self in spontaneous thoughts

In the first experiment (Babo-Rebelo et al. 2016a), with MEG, we studied the self in spontaneous thoughts, by distinguishing between the "I" and the "Me". The amplitude of heartbeat-evoked responses (HERs) in the posterior cingulate / ventral precuneus regions (PCC/vPrc) co-varied with the engagement of the "I" in the ongoing thought (Figure 10). The "Me" dimension was associated with HERs in the ventromedial prefrontal cortex (vmPFC). We further showed that these results were specific to each self-dimension. We then replicated these results with intracranial recordings, by showing a trial-by-trial parametrical modulation of HERs along with the levels of involvement of each self-dimension in thoughts (Babo-Rebelo et al. 2016b). The main results were found in midline regions of the default-network (DN), but a region of interest analysis (ROI) additionally revealed that HERs in the right anterior insula (rAI) were also modulated by the degree of engagement of the "I" in thoughts. We argued that the neural monitoring function of the DN could be related to self-processing, and that this integration could generate a subject-centered reference frame from which selfhood can emerge. Our initial hypothesis focused mostly on the "I" and we had no specific hypothesis for the "Me". Our results show that both aspects of the self are associated with HERs, but in different regions.

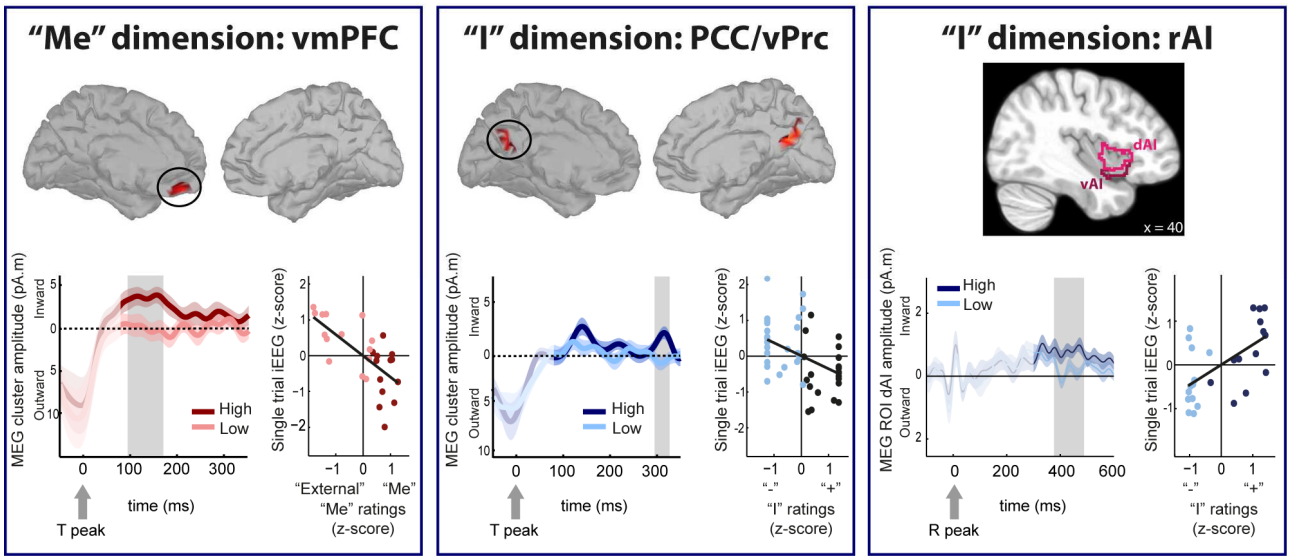


Figure 10: Main MEG and iEEG results of the experiment on spontaneous thoughts.

2. *HERs distinguish self- and other-imagination*

The first study was based on spontaneous thoughts, which content by definition cannot be controlled by the experimenter. In a follow-up study, we wanted to orient participants' thoughts, as well as contrast self and other. In the literature on the self, this contrast is more common than the gradual levels of self that we implemented in the first study. We therefore performed another MEG experiment, where participants had to imagine themselves (from the first-person perspective) or a friend (from the third-person perspective).

In a whole-brain analysis, we found differential HERs in medial motor regions (anterior precuneus, mid-cingulate and supplementary motor area) and, using a region of interest analysis, in the vmPFC (Figure 11). Moreover, the amplitude of the effect for Self at the sensor level was positively correlated with the propensity of participants to daydream in their daily lives. The amplitude of the effect for Other was in turn negatively correlated with ratings on the vividness of the perspective. While PCC/vPrc and rAI showed significant differences in HERs in the first experiment, here we did not find any significant differences in PCC/vPrc nor in the rAI. Future analyses will aim at further exploring the trial-by-trial ratings on the vividness of the adopted perspective.

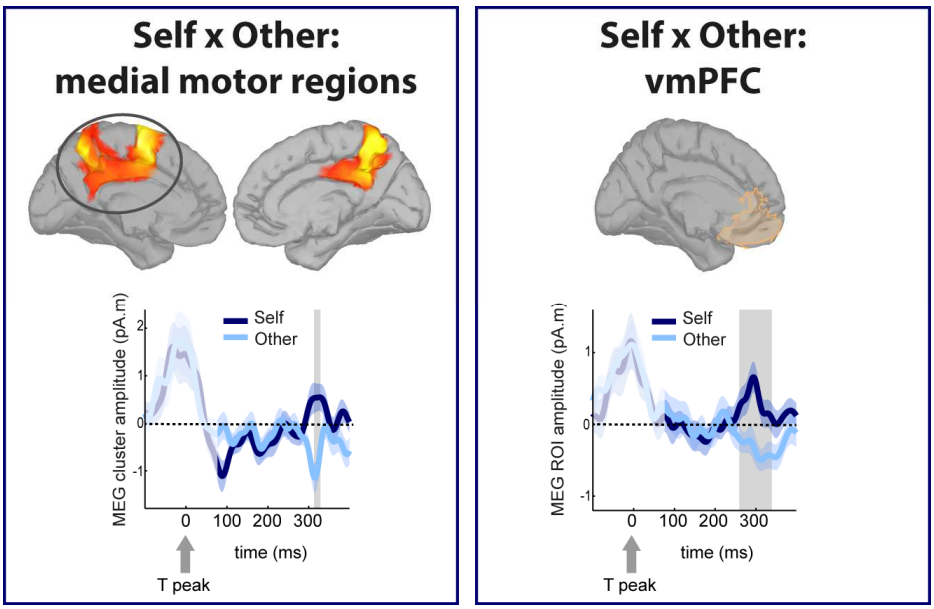


Figure 11: Main MEG results on the experiment on imagination.

B. What do these results tell us about the self?

1. The "I" and the "Me": two distinct and graded dimensions of the self in spontaneous thoughts

Following the criticisms of Christoff and colleagues on the neuroscience of the self (Christoff et al. 2011) which we exposed in the Introduction, we here aimed at studying the more implicit form of self, the "I". Our results suggest that this dimension exists, and that it is different from the "Me" dimension (Babo-Rebelo et al. 2016a). This dissociation occurs at the brain level, despite the strong correlation between the two at the behavior level. In the iEEG experiment (Babo-Rebelo et al. 2016b), this dissociation was less clear, which could be explained by the low number of trials or a less accurate comprehension of the meaning of the scales by the patients who went through a shorter version of the training on the scales.

In the second paper (Babo-Rebelo et al. 2016b), we showed that HER amplitude was parametrically modulated by the continuous ratings on each self-related scale. This further shows that in spontaneous thoughts the self is expressed in a continuum rather than an all-or-none fashion. Continuous ratings allow a more precise and intuitive classification of complex thoughts and allow expressing the vividness of the first-person perspective or the intensity of the introspective thought.

The operational distinction that we made between the "I" and the "Me" relates to the phenomenological distinction between the self-as-subject and the self-as-object (Table 1 - Introduction). However, the "I" for us was verbally expressed in thoughts and could be referred to not only in the present, but also in the past and future. These are two major differences with the philosophical literature in the domain, which considers this form of self as being the immediate, implicit and non-verbal experiential self. Here, we adopted a more liberal perspective on the "I", by associating it with first-person perspective *in the thought*.

2. What is contrasted when we compare Self and Other?

In the imagination task, Self and Other conditions differed for at least three reasons: 1) the person being imagined (self or friend), 2) the perspective (first-person or third-person), 3) the type of mental imagery (action simulation or action observation). Imagining the self from the first-person perspective, in simulation, is

the best way to approach the idea of the self *as the subject*. Finding an appropriate contrast condition with Other is more challenging. The subject dimension would probably still be present if the friend is imagined from the first-person perspective. In addition, such an experimental condition would probably be unnatural for participants. Imagining the friend from a third-person perspective seemed a better option, as also performed by Ruby and Decety (Ruby & Decety 2001). Self and Other are thus intrinsically different in other respects than purely the person being imagined. This might be true in a number of paradigms of the literature on the self.

Using a Self vs Other contrast emphasizes the "Me" dimension (Legrand & Ruby 2009), because thoughts can be self-directed or directed to someone else. Addressing the "I" with this contrast is more debatable, because even if I am imagining my friend, I am still present as the observer, as the "I". This could explain why we do find the vmPFC in the imagination task but not the PCC/vPrc.

C. Consistency of the results between tasks

In the spontaneous thoughts experiment, we found that the PCC/vPrc was associated with the "I" and the vmPFC with the "Me". In the imagination task, we do not observe any effect in the PCC/vPrc (Table 2). This is surprising because one would expect that the "I" is more engaged when imagining oneself, than when imagining someone else. Conversely, we observe an effect in the vmPFC (Table 2), suggesting that a reflective aspect of the self is engaged. It is worth underlying here that the latency of the HER effects in the vmPFC was quite different between tasks (Table 2). The Self-Other difference in the imagination task took place around 300ms after the T-peak, whereas in the spontaneous thoughts task the effects appeared much earlier (around 150ms after the T-peak). In the imagination task, the regions of the precuneus and cingulate cortex were clearly more anterior than the PCC/vPrc of the spontaneous thoughts task (Figure 12). Maybe the Self vs Other contrast is not the best way to assess the "I", and looking at different levels of engagement of the first-person perspective (maybe by exploring the ratings on the vividness of the perspective) during imagination would better correspond to the "I" dimension.

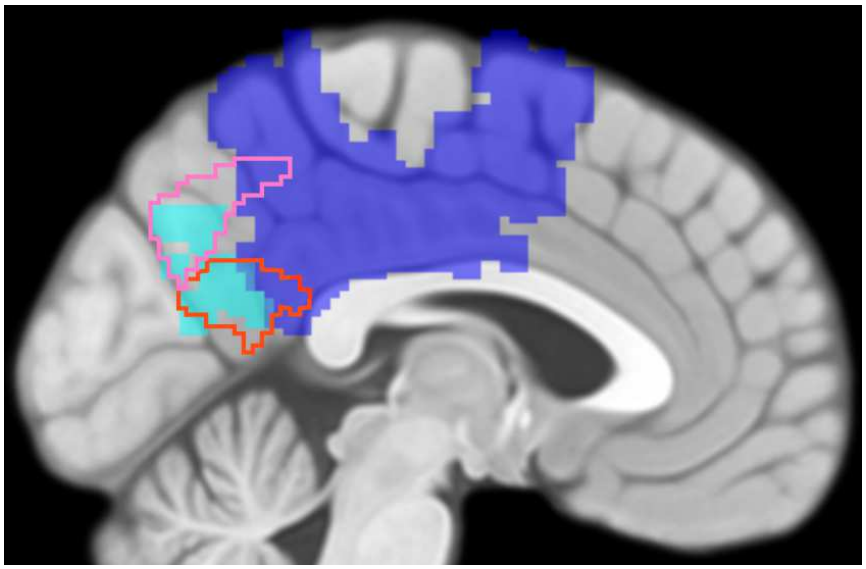


Figure 12: Self-related HERs in mid- and posterior medial cortices. Regions in blue showed differential HERs during Self- vs Other-imagination. Regions in cyan showed self-related HERs during spontaneous thoughts. Pink and red outlines correspond respectively to the ventral precuneus, and ventral posterior cingulate cortex (Bzdok et al. 2015)**. Sagittal view, x=-3.**

The effect in the rAI in the spontaneous thoughts task was less robust than the effects in midline DN regions, so the fact it is absent in the imagination task is not surprising. Even though this region is often considered as the primary viscerosensory region, and has been hypothesized to "engender human awareness" (Craig 2009), it seems that its role in the generation of self-related HERs is marginal.

The involvement of midline motor regions in the imagination task is coherent with the fact that half of the scenarios were explicitly cueing actions. Moreover, this result is consistent with the full-body illusion experiment showing a modulation of HERs depending on the intensity of the illusion (Park et al. 2016). This suggests that the imagination task engages a strong bodily dimension. It is noteworthy that midline motor regions were also exhibiting differential HERs along with the "I" dimension in the spontaneous thought experiment, at the same latency as in the imagination experiment (uncorrected results, Table 2, Figure 13). Possibly, some spontaneous thoughts were related to action simulation. It seems thus that some of the differences between the two tasks could be driven by a different balance in the content of thoughts (more or less mental imagery, self-reflection…).

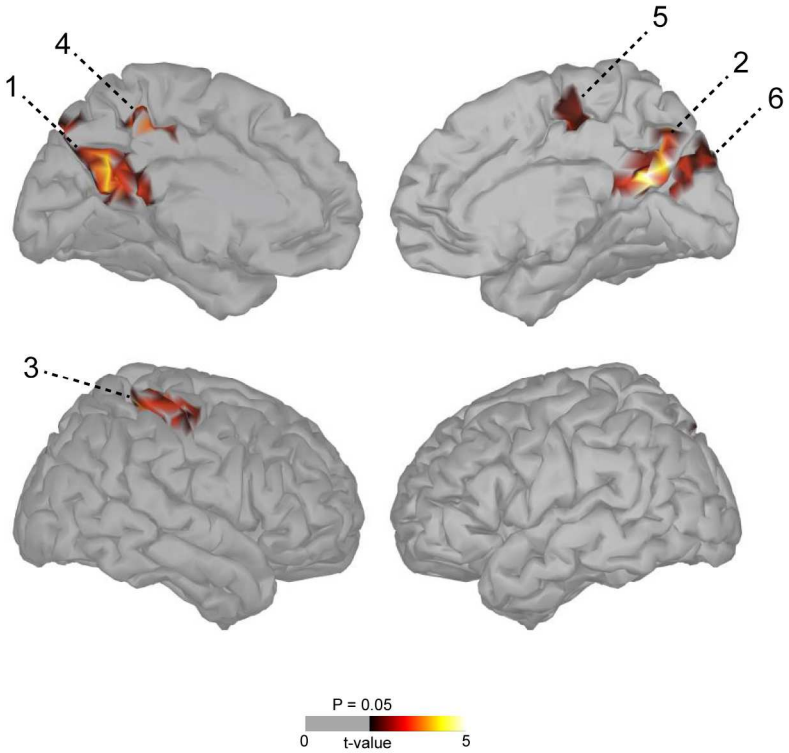


Figure 13: Regions responding differentially to heartbeats along the "I" scale in the spontaneous thoughts experiment, with a more liberal threshold (>20 contiguous vertices at uncorrected p<0.05). 1: Left precuneus, posterior cingulate. 2: Right precuneus, posterior cingulate. 3: Postcentral gyrus. 4: Left precuneus, mid-cingulate. 5: Right supplementary motor area. 6: Right cuneus, calcarine. (AAL atlas) The regions referenced with number 1 are the only ones that survive correction for multiple comparisons and correspond to those presented in the paper.

In the spontaneous thought task, only high levels of involvement of the self in thoughts led to a significant increase of brain activity relative to zero. In the imagination task, both Self and Other conditions elicited a change in brain activity. The reason for this difference is still unclear, but it could be related to the fact that in one case we are contrasting high and low levels of self-relatedness, and in the other case we are contrasting Self with Other.

	vmPFC	PCC/vPrc	rAI	Motor medial regions
Spontaneous thoughts task	✓ 150ms	✓ 300ms	**(✓)** 180ms	✓ ~300ms
Imagination task	**(✓)** 300ms	-	-	✓ 300ms

Table 2: Summary of the main results. The ✓ symbol indicates that a significant effect was found in the corresponding structure and task, in a whole-brain analysis. (✓) indicates that the corresponding effect was found in a region-of-interest analysis. In grey, the effect was uncorrected for multiple comparisons. Below the symbol, we indicate the timing around which the HER effect takes place (relative to the T-peak).

D. What do these results tell us about spontaneous vs oriented thoughts?

A task relying uniquely on spontaneous thoughts has some inherent drawbacks, in particular because thoughts are private and can only be assessed with self-report (Hurlburt & Heavey 2001, Smallwood & Schooler 2015). To try and overcome these issues, we developed a thorough training and testing procedure and allowed participants to skip the responses if they were unsure. Performing a task where the only instruction is to mind wander is quite hard to perform for cognitive neuroscientists, but actually very easy for participants who report having no problems in letting their minds wander, as Hurlburt and Heavey also observed (Hurlburt & Heavey 2001).

In the imagination task, we gave instructions for participants to orient their thoughts. We saw that our effects were not totally stable during the 6 seconds of the imagination period, suggesting that there might be some effort associated with maintaining imagination (and possibly suppressing task-irrelevant mind wandering). In the imagination task, we found the involvement of the vmPFC, which is part of the DN. This suggests that the DN is not exclusively related to *spontaneous* cognition, an idea that has been discussed lately (Christoff et al. 2016).

E. Proposal of a mechanism for the implementation of the self

1. What is this signal?

MEG sensors can directly pick up the electrical activity from the heart. It could be argued that HERs are not of neural origin, but result from differences in cardiac activity. We made different methodological choices to try and rule out this possibility (ICA correction, choice of the time window of analysis, cardiac parameter analyses). The fact that we replicate the MEG results with iEEG, which is less affected by the cardiac-field artifacts (Kern et al. 2013), is a strong argument to believe we are truly looking at neural signals.

Yet, the characteristics of these neural evoked responses are unclear. HERs do not have a clear topography or clear components. HERs have been found in different regions, even in regions that are not primary targets of visceral signals. How these

signals can be relayed to these structures has not been explored yet. The latencies of HER effects seem also quite variable, as we found early and late effects, even in the same structures (the vmPFC). How HER latencies are affected by changes in heart rate (during physical exercise, for instance) is still unknown. Besides, what we should consider as the time-locking point is unclear. We chose to compute HERs locked to the T-peak because it is closer to the artifact-free time window (Dirlich et al. 1998) that we analyzed, but other groups have computed HERs locked to the R-peak. This question is not merely methodological, it is conceptual as well. Since there is a time jitter between the R- and T-peaks, what is the event that is actually causing these neural evoked responses?

This raises the question of the origin of HERs. We exposed in the Introduction the hypothesis that they are originated by the discharge of mechanoreceptors in the heart wall and aortic arch at each heartbeat. This is a plausible origin for HERs, but it remains an open question. Recently, it was shown that changes in blood pressure in the mouse brain lead to changes in the firing activity of neurons (Jung Kim et al. 2016). Vasculo-neuronal coupling could thus be another mechanism generating HERs.

2. Three hypotheses to explain the link between HERs and the self

Our results show a *correlation* between HER amplitude and the self. What mechanism can underlie this correlation and what can it mean?

A first hypothesis is that HERs are a byproduct of brain activity, meaning that a region that is particularly active will show changes in brain activity following a heartbeat. Therefore, if two conditions induce different levels of activity in a certain region, the changes in activity following a heartbeat will consequently differ as well. This idea would be supported by the fact that differential HERs are systematically found in regions which are expected to be differently activated: default-network regions during the resting state (Babo-Rebelo et al. 2016a,b; Park et al. 2014), bodily/motor regions during full-body illusions and action imagery (Park et al. 2016)(Babo-Rebelo & Tallon-Baudry, *in prep*.), and viscerosensory regions during attention to heartbeats (Canales-Johnson et al. 2015, Pollatos et al. 2005a). More intriguingly, we find HERs in regions which are not particularly known to be targets

of ascending visceral signals (PCC/vPrc). However, using the surrogate heartbeats test, we and others (Park et al. 2014, 2016) have shown that HER differences cannot be trivially explained by different baseline activities or slow-fluctuations of brain activity. Still, brain activity could differ in other ways that we are not able to easily measure. Another possibility would be that active regions are more responsive to heartbeats. If we push the argument even further, active regions could be more responsive to any kind of stimulus, whether a heartbeat or any other stimulus. This would require a difference in cortical excitability, which does not seem to be the case since we did not observe differences in alpha power (Babo-Rebelo et al. 2016a, Park et al. 2014)(Babo-Rebelo & Tallon-Baudry, *in prep*.). Still, one should keep in mind that not observing a difference does not mean that a difference is not there.

A second hypothesis would be that ascending visceral signals define the self. In a spontaneous thought, the cardiac signal itself would indicate that the thought is self-related or not. However, this would imply that cardiac signals carry information about the self, in accordance with the ongoing thought. How would the cardiovascular system be informed about the contents of our thoughts? This would require particular brain-body loops that would modify bodily signals. Importantly, in our experiments, we did not observe any differences in the cardiac parameters that we were able to measure (heartrate, heartrate variability, and blood pressure in (Park et al. 2014)), but more subtle or different parameters could vary. Furthermore, this hypothesis implies that the self would be implemented at each cardiac cycle, which is hard to reconcile with our experience of continuity.

A third hypothesis would be that of an interaction between brain activity and ascending visceral signals which would causally implement the self. Cardiac (as well as other visceral) signals would function as a ticking clock, periodically sending signals to the brain, indicating that a body is there. These signals could interact with ongoing brain activity, in a way that would generate larger or smaller responses to heartbeats. The amplitude of these responses would be a marker of the self. Cardiac signals would contribute to a body-centered reference frame, which would be used by the brain to anchor thoughts to the self. Bodily signals do not have to carry information, their existence is enough. Note that given the wide availability of visceral signals, this mechanism could take place in many different regions of the brain, wherever a self vs non-self distinction is relevant for the task at play.

This proposal still arises many questions. What kind of mechanism could generate the interaction between cardiac signals and brain activity? Are self-related HERs larger in amplitude than non-self-related HERs? Is it truly a common code throughout the brain? Understanding what kind of mechanism can generate such an interaction is a difficult question, which would necessitate the experimentally challenging step of moving from correlation to causation. The cases of patients with artificial hearts could be potentially interesting. Our guess is that a beating artificial heart is likely to provide the cyclic signal that this mechanism requires. Recently, new artificial hearts are being developed, which function with a continuous flow. What would happen in this case, where the heart does not beat anymore? That remains an open question.

IX. Appendix

A. Article I: Phase-amplitude coupling at the organism level: the amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm

NeuroImage 146 (2017) 951–958

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Phase-amplitude coupling at the organism level: The amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm

Craig G. Richter [a,b,*], Mariana Babo-Rebelo [a], Denis Schwartz [c], Catherine Tallon-Baudry [a,*]

[a] Laboratoire de Neurosciences Cognitives (ENS – INSERM), Ecole Normale Supérieure - PSL Research University, Paris, France
[b] Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Frankfurt, Germany
[c] Sorbonne Universités, Inserm U 1127, CNRS UMR 7225, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle épinière, ICM, Paris, France

ARTICLE INFO

Article history:
Received 24 June 2016
Accepted 20 August 2016
Available online 21 August 2016

Keywords:
Cross-frequency coupling
Phase-amplitude coupling
Resting-state dynamics
Alpha rhythm
Gastric rhythm
Brain-viscera interactions

ABSTRACT

A fundamental feature of the temporal organization of neural activity is phase-amplitude coupling between brain rhythms at different frequencies, where the amplitude of a higher frequency varies according to the phase of a lower frequency. Here, we show that this rule extends to brain-organ interactions. We measured both the infra-slow (~0.05 Hz) rhythm intrinsically generated by the stomach – the gastric basal rhythm – using electrogastrography, and spontaneous brain dynamics with magnetoencephalography during resting-state with eyes open. We found significant phase-amplitude coupling between the infra-slow gastric phase and the amplitude of the cortical alpha rhythm (10–11 Hz), with gastric phase accounting for 8% of the variance of alpha rhythm amplitude fluctuations. Gastric-alpha coupling was localized to the right anterior insula, and bilaterally to occipito-parietal regions. Transfer entropy, a measure of directionality of information transfer, indicates that gastric-alpha coupling is due to an ascending influence from the stomach to both the right anterior insula and occipito-parietal regions. Our results show that phase-amplitude coupling so far only observed within the brain extends to brain-viscera interactions. They further reveal that the temporal structure of spontaneous brain activity depends not only on neuron and network properties endogenous to the brain, but also on the slow electrical rhythm generated by the stomach.

Introduction

Phase-amplitude coupling (PAC) is a fundamental organizational rule where the amplitude of a high-frequency oscillation varies according to the phase of a lower frequency oscillation (Bragin et al., 1995; Canolty et al., 2006; Schroeder and Lakatos, 2009; Buzsaki, 2010). This rule has been recently shown to also govern the temporal organization of spontaneous large-scale brain activity in humans (Osipova et al., 2008; Roux et al., 2013; Florin and Baillet, 2015; Weaver et al., 2016). Here, we propose to extend the hierarchical organization of PAC to brain-viscera interactions. The brain at rest is not a closed system as it constantly receives information from visceral organs (Mayer, 2011; Critchley and Harrison, 2013; Furness et al., 2013). Some organs may provide an external source of slow frequency rhythms relayed to the brain and contributing to the temporal organization of resting-state brain dynamics. Specifically, we hypothesize that the gastric basal rhythm, an infra-slow electrical oscillation intrinsically and continuously generated by the stomach, may influence resting-state brain dynamics.

The stomach contains a specific cell type - the interstitial cells of Cajal (Sanders et al., 2006; Sanders et al., 2014) - at the interface between the enteric nervous system and gastric smooth muscles, that intrinsically generate an electrical slow wave at ~0.05 Hz (3 cycles per minute). During digestion, the gastric basal rhythm sets the pace of muscle contraction, but the rhythm is generated at all times, even in the absence of contraction (Bozler, 1945), or when the stomach is experimentally disconnected from the central nervous system (Suzuki et al., 1986). Gastric interstitial cells of Cajal form synapse-like connections with afferent sensory neurons (Powley and Phillips, 2011) that, via spinal and vagal nerve pathways and various subcortical relays, target a number of cortical structures comprising notably the insula, ventral anterior cingulate cortex and somatosensory cortex (Ito, 2002; Mayer, 2011; Critchley and Harrison, 2013; Furness et al., 2013). The stomach may thus be considered as an autonomous electrical pacemaker that may continuously feed the brain with a slow oscillatory input

* Corresponding authors.
E-mail addresses: craiggrichter@gmail.com (C.G. Richter), catherine.tallon-baudry@ens.fr (C. Tallon-Baudry).

http://dx.doi.org/10.1016/j.neuroimage.2016.08.043

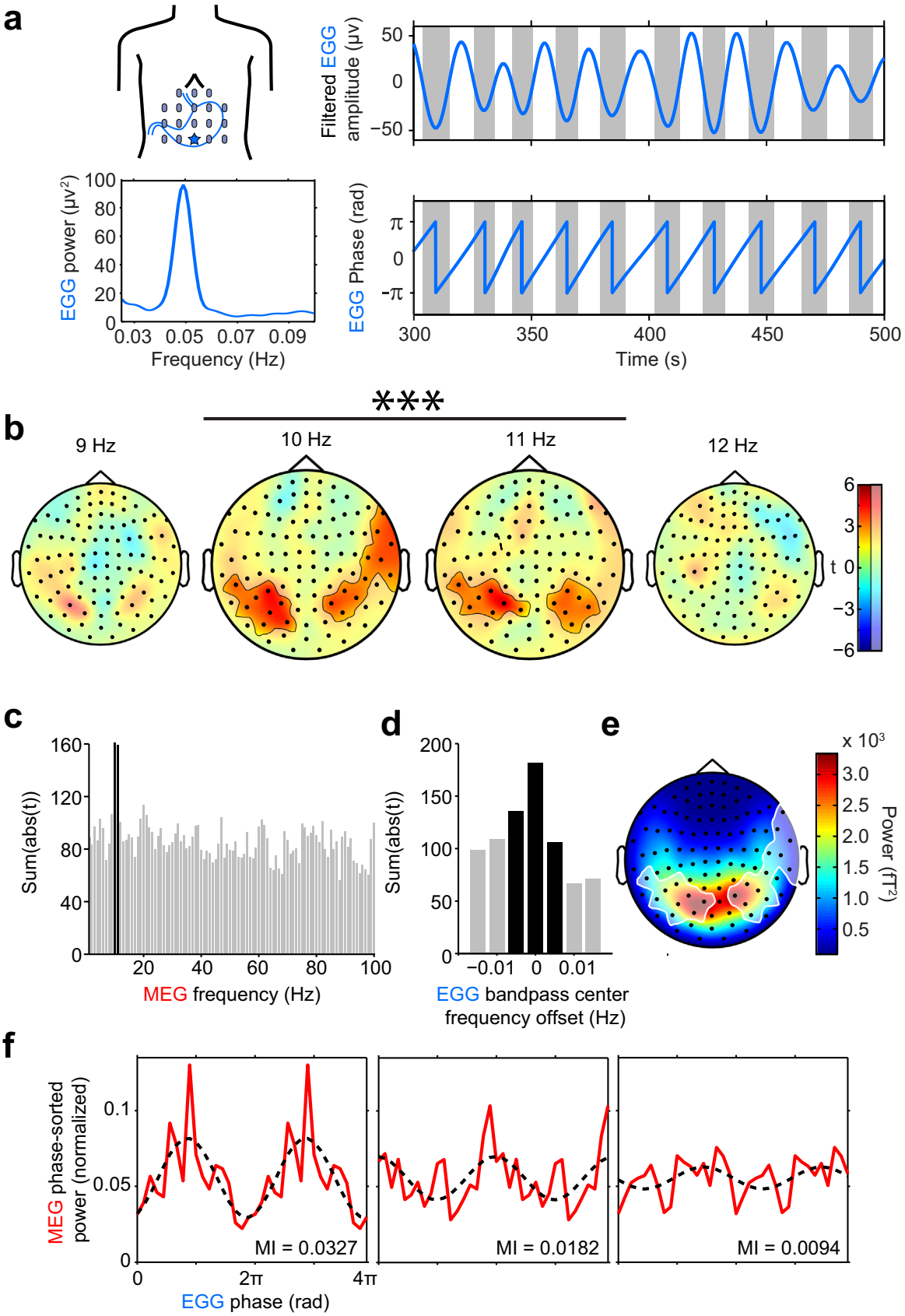


Fig. 1. Gastric-alpha coupling. (a) Electrogastrogram (EGG) recording in a single participant. Left, electrode montage with the star indicating the electrode with the largest EGG amplitude and corresponding EGG power spectrum below. Right: 200 s filtered EGG signal and corresponding phase below. (b) Statistical maps of gastric phase - MEG power coupling at different frequencies. Significant gastric-MEG coupling occurred only at 10 and 11 Hz, in the two clusters indicated by black outlines and saturated colors (Monte-Carlo $p=0.0008$ for both clusters, corrected). (c) Summary statistics of gastric coupling strength across all sensors, for brain frequencies between 1 and 100 Hz, display a sharp peak at 10 and 11 Hz. (d) Coupling is specific to gastric frequency: summary statistics of EGG-alpha coupling strength across all sensors decreases when offsetting the filter above or below EGG peak frequency, that is presented at 0. Black bars in (c) and (d) indicate significant coupling. (**e**) Topographical map of 10–11 Hz power, grand average across participants. The clusters of significant gastric-alpha coupling at 10–11 Hz are overlaid in white. (f) Three examples of phase-amplitude coupling profiles, in the participant with the largest (left), median (middle) and smallest (right) MI. Profiles are presented over two gastric cycles (4π) for clarity. MEG average power in each bin was normalized by the sum of the average power across bins. The dashed black line is a cosine fit that emphasizes the 1:1 coupling between alpha power and gastric phase.

constraining resting-state brain dynamics.

We tested whether the phase of slow gastric oscillations is coupled to the amplitude of higher-frequency brain rhythms at rest in humans. We recorded brain activity with magneto-encephalography (MEG) along with gastric electrical activity measured from cutaneous electrodes placed on the abdomen (Fig. 1A), a technique called electrogastrography (EGG) (Koch and Stern, 2004), from 17 participants at rest with eyes open for 12 min.

Material and methods

Participants

Seventeen right-handed adult participants (mean ± sem age: 23.9 ± 0.62, range 20–29; 8 males; mean body-mass index 22.02 ± 0.62, range 17.5–26.1) with normal or corrected to normal vision took part in the study. None of the participants had any previous history of neurological, psychiatric or digestive disease. Participants had been fasting for at least 2 hours before the recordings. They signed a written informed consent and were paid for participation. All procedures were approved by the Ethics Committee CPP Ile de France III and were in accordance with the Helsinki declaration.

Procedure and recordings

Participants fixated a central black fixation mark (black dot, radius 0.13° of visual angle, surrounded by a black circle, radius 0.38° of visual angle) presented on a gray background at a viewing distance of 80 cm for 12 min. Participants were instructed to stay still, to fixate the central mark and to let their mind wander, avoiding any structured strategy such as counting or mentally reciting a text. Continuous magneto-encephalographic (MEG) signals were collected using a whole-head MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX MEG system) at a sampling rate of 1000 Hz and online low-pass filtered at 330 Hz. The electrogastrogram (EGG) was recorded via 19 disposable cutaneous electrodes (17 active, 1 reference and 1 ground) placed on the abdomen and acquired simultaneously with MEG data (DC recordings, low-pass filter at 330 Hz). In classical EGG montages, the reference electrode is located in the upper right part of the abdomen, and active electrodes are placed over the left part of the abdomen (Chen et al., 1999), where the stomach lays. We extended this montage to create a bilateral grid of EGG electrodes placed over four regularly spaced rows (Fig. 1a). In each participant, we first determined the midpoint between the xyphoid process and the umbilicus. The central electrode of the second row was located 2 cm above this midpoint (Chen et al., 1999). The vertical position of the top-row was then determined as the intersection of a 45° line originating from the central electrode of the second row, and the left mid-clavicular line. The horizontal positions of rows 3 and 4 were distributed such that the vertical spacing between each row was equal. The electrodes were horizontally centered on the midline and were evenly distributed between the left and right mid-clavicular lines. The first row consisted of 3 electrodes, with the rightmost electrode being used as a reference. The subsequent rows consisted of 5 electrodes. The ground electrode was located on the participant's left costal margin. An electrocardiogram (ECG) was simultaneously recorded. Eye position and pupil diameter were monitored with an EyeLink 1000 (SR Research) and simultaneously recorded with MEG, EGG and ECG data.

MEG data preprocessing

Signal Space Separation (tSSS) was performed using MaxFilter (Elekta Neuromag) to remove external noise. Subsequent analysis was conducted on magnetometer signals. The cardiac artifact was corrected using Independent Component Analysis (ICA), as implemented in the FieldTrip toolbox (Oostenveld et al., 2011). Briefly, the 12 min. resting period was divided into 5 s segments to compute ICA components. The number of independent components to be identified was the rank of the time x trial matrix. The continuous magnetometer data were then decomposed according to identified ICA components. The ICA-decomposed MEG signals and ECG data were epoched from 250 ms before to 400 ms after each R-peak and the pairwise phase-consistency (PPC) (Vinck et al., 2010) was computed between the ICA-decomposed signals and the ECG signal to isolate those components most reflective of ECG activity. Components with large PPC values and topographies matching the stereotypical ECG artifact were rejected from the continuous MEG data (mean 1.71 ± 0.41 sem components rejected). Blink artifacts were defined as the blink intervals identified by the EyeLink eye-tracker system padded by ± 100 ms. On average, 8.92% ± 2.19 sem of the total recording time was marked as contaminated by blink artifacts and was excluded from the analysis. ICA-corrected magnetometer data were then downsampled to 400 Hz and submitted to a Hann tapered 1 s window FFT, computed from 0 to 720 s at 0.050 s steps. The squared-magnitude of the resulting complex Fourier coefficients was used to generate the power envelope time series with a 20 Hz sampling rate for frequencies ranging from 1 to 100 Hz in 1 Hz steps.

EGG processing

EGG power at each abdominal electrode was computed via a Hann tapered FFT, using Welch's method with a 200 s window moving in 50 s steps. For each participant, the electrode exhibiting the largest spectral peak in the 0.05 ± 0.01 Hz range, centered in the normogastric range (Riezzo et al., 2013), was selected for further analysis. To identify and mark EGG artifact periods, the raw signal was filtered between 0.01 and 0.5 Hz to isolate EGG related variance, and the standard deviation was computed over the trial. Segments of this filtered signal exceeding 4 standard deviations were marked as artifacts. These periods were padded by ± the filter order used to isolate the EGG peak frequency (± 0.02 Hz of the peak frequency, see below) to compensate for temporal smearing of the artifact by the filter. On average 12.85% ± 2.71 sem of the total recording was discarded due to presence of artifacts in the EGG signal, mostly due to participant movement. The raw EGG was then downsampled to 20 Hz and filtered using a frequency sampling designed finite impulse response filter (Matlab: FIR2), with a bandwidth of ± 0.02 Hz of the peak EGG frequency, and a transition width between the passband and stopband of 15% of the upper and lower passband frequencies. The filter order was determined as the number of samples corresponding to 3 cycles of the lower passband frequency. Importantly, filter width was large enough to capture slower and faster gastric episodes. Filtered data thus retained all the frequency variability intrinsic to the gastric rhythm necessary for the statistical procedure we used (see below). The filter width was sufficiently narrow enough to exclude any contribution from respiration. The filtered EGG signal was then Hilbert transformed and the analytic phase was derived.

Phase-amplitude coupling (PAC)

As a result of the preprocessing steps, we obtained a pair ($\phi_{gastric}(t)$, $Pow_{MEG}(t)$), where $\phi_{gastric}$ is the phase of the gastric rhythm, and Pow_{MEG} the power of the MEG signal in a given

frequency band at a given sensor, at each sample t of the artifact-free epochs. The EGG phases were sorted into 18 bins spanning the [–pi, pi] interval, and corresponding MEG power was averaged for each phase bin. MEG power sorted by EGG phase bin defined the PAC profile (Fig. 1f). To quantify the deviation of the PAC profile from a uniform distribution, we computed the modulation index (MI) (Tort et al., 2010). Briefly, when the MEG power shows no systematic relationship to the EGG phase, MEG power in each EGG phase bin will tend toward the overall average MEG power, resulting in a flat, or uniform, distribution. The MI of Tort et al. (2010) specifically measures deviation from a uniform distribution, and thus in this case a correspondingly low MI value will result. Alternatively, if the MEG signal power systematically differs across EGG phase bins, the PAC profile will deviate from a uniform distribution and MI will be larger. As shown by (Tort et al., 2010), MI is sensitive to 1:1 coupling but also to higher 1:m coupling modes (Palva et al., 2005).

Statistical determination of significant clusters of PAC

The statistical determination of significant clusters of phase-amplitude coupling was a two-step process. We first estimated, for each participant, chance-level PAC at each sensor and frequency. We then determined, at the group level, sensors and frequency where a significant difference between observed coupling and chance-level coupling differed. Those steps are detailed below.

We first estimated the level of PAC expected by chance and the corresponding chance-level MI for each participant, magnetometer and MEG frequency. We created surrogate data where the relationship between EGG and MEG signals was disrupted by shifting EGG phase and MEG power signals relative to one another by a random time interval exceeding $\pm$ 60 s, i.e. about 3 gastric cycles. Data at the end of the record were wrapped to the beginning, as in the cutting/swapping procedure proposed by Bahramisharif et al. (2013). This procedure best preserves phase autocorrelation and is much more conservative than the random shuffling of the full time series (Weaver et al., 2016). In other words, from the original pairs ($\phi_{\mathrm{gastric}}(t)$, $\mathrm{Pow}_{\mathrm{MEG}}(t)$) we created surrogate pairs ($\phi_{\mathrm{gastric}}(\mathrm{t})$, $\mathrm{Pow}_{\mathrm{MEG}}(t+\tau)$) where τ is a value randomly chosen between 1 and 11 min. Because the filtered EGG signal, and MEG power envelope are not pure sine waves, but physiological signals that exhibits spontaneous increases and decreases in frequency, any link between gastric phase and brain rhythms is disrupted in the surrogate data. For each participant, MEG sensor and frequency, we obtained a distribution of surrogate MI values by creating 1000 surrogate data sets, corresponding to 1000 random τ, and computing the associated MIs. We defined the chance level, for each participant, sensor and MEG frequency, as the median of surrogate MI values.

We tested whether the empirical MI significantly differed from chance level MI at the group level using a cluster-based permutation procedure (Maris and Oostenveld, 2007), as implemented in FieldTrip (Oostenveld et al., 2011), that extracts significant differences between two conditions, across sensors and MEG frequencies, while intrinsically correcting for multiple comparisons. Briefly, this procedure entails comparing empirical MI with the corresponding chance level MI value across participants using a *t* test at each sensor and frequency. Candidate clusters are defined in space as sensors exceeding the first level t-threshold ($p < 0.05$, two-sided) and that are connected to at least 2 neighboring sensors that also exceed this threshold, and across adjacent frequencies that exceeded the first level t-threshold. Each candidate cluster is characterized by a summary statistic corresponding to the sum of the *t*-values across the sensors and frequencies defining the cluster. The second-level statistic, i.e. whether a given sum of *t*-values in the candidate cluster could be obtained by chance, was determined by computing the distribution of cluster statistics under the null hypothesis. In practice, we randomly shuffled the labels 'empirical' and 'chance' 10,000 times, applied the clustering procedure and retained the largest positive and negative clusters from each permutation. Across the 10,000 permutations one can thus build the distribution of cluster statistics under the null hypothesis, which is then used to assess the empirical clusters for significance. Because the largest positive and negative clusters are retained at each permutation, this method intrinsically controls for multiple comparisons over sensors and frequencies (Maris and Oostenveld, 2007). The resulting clusters are described by their summary statistics, corresponding to the sum of t-values for each time sample and sensor belonging to the cluster, and by their MonteCarlo *p*-value describing significance at the cluster level corrected for multiple comparisons across sensors and frequencies.

Explained variance

To determine the percentage of fluctuations of brain activity at a given frequency explained by the phase of the gastric rhythm, we computed the ratio between the variance of the original MEG amplitude envelope and the variance of the PAC profile, i.e. the MEG amplitude envelope sorted by gastric phase (Fig. 2). When MEG amplitude shows no systematic variation with EGG phase, then the distribution of phase-sorted amplitude will approach uniform, which will yield a low variance computed across the bins. Alternatively, when the MEG data is systematically modulated by EGG phase, the distribution of phase-sorted amplitude will be non-uniform giving rise to a larger variance across bins. The variance across bins of the phase-sorted amplitude is divided by the variance of the original non-phase organized signal, which gives the proportion of the MEG amplitude fluctuations in the original signal that is explained by EGG phase, or, in other words, the explained variance. In practice, we computed the ratio between the variance of the time-varying MEG 10–11 Hz amplitude envelope binned by EGG phase and the variance of the original MEG 10–11 Hz amplitude envelope smoothed in time-windows of a duration equal to the length of one phase bin, using a zero-phase moving average filter.

Source analysis

We used a beamformer-based source localization technique to obtain a time series of 10–11 Hz power per voxel, per participant. A 5 mm grid spanning the MNI ICBM 152 nonlinear high-resolution (0.5 mm) template brain was constructed. This grid was warped to the anatomy of each participant based on his or her individual MRI. The ICA corrected magnetometer signals were downsampled to 50 Hz, and zero-phase filtered between 10 and 11 Hz (FIR frequency sampling filter, transition band of 15% of the upper and lower passbands, order=200). A spatial filter was constructed using an LCMV beamformer and a single-shell head model, implemented in the FieldTrip toolbox (Oostenveld et al., 2011). Data containing blinks were excluded from spatial filter construction. Since resting-state data cannot be contrasted to another condition, the leadfields were normalized using the default parameter of 0.5 to reduce power bias towards the center of the head. Source time series of the 10–11 Hz data were constructed by projecting the 12 min 10–11 Hz filtered data segment through the spatial filter and taking the magnitude of each dipole along its principal axis. The power envelope was then determined as the squared-magnitude of the Hilbert transform. The resulting power envelope was downsampled to 20 Hz. These virtual source time series were used to determine PAC values using the same computation as at the sensor level. A one-tailed cluster statistic was then computed, using the same surrogate data sets and clustering procedure as

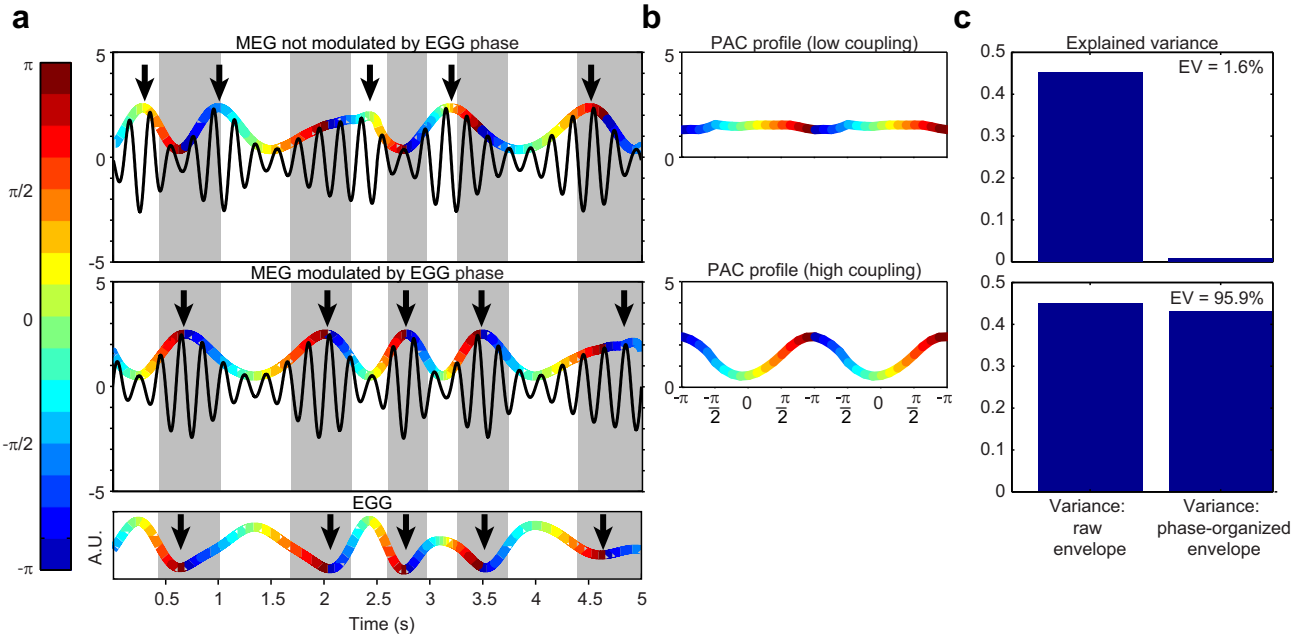


Fig. 2. Rationale for the quantification of explained variance. (a) *Bottom*, simulated EGG signal, varying over time in amplitude (arbitrary units) and phase (color code). Arrows indicate EGG troughs. *Middle*, simulated rhythmic MEG activity with an amplitude envelope modulated by the phase of the EGG signal. Arrows indicate peaks in the MEG amplitude envelope. The color code in the amplitude envelope refers to EGG phase. Both arrows and color-coded phase reveal the systematic link between the MEG amplitude envelope and EGG phase. *Top*, simulated MEG signal with an amplitude envelope that varies over time but that is not coupled to EGG phase. Arrows indicate peaks in the MEG amplitude envelope. (b) Phase-amplitude profiles, i.e. MEG amplitude envelope sorted by EGG phase. When MEG amplitude has little relationship with EGG phase (top row), the phase-amplitude profile is flat. When alpha amplitude is tightly linked to gastric phase (bottom row), the phase-amplitude profile shows modulations as large as in the original non phase-sorted data. (c) Variance of the raw envelope as presented in (a) and variance of the phase-sorted envelope as presented in (b). Explained variance (EV) is the ratio between the variance of the phase-organized envelope and the variance of the raw envelope, expressed in percentage.

described at the sensor level, to determine regions showing significant gastric-brain coupling at the source level. This test utilized a first level threshold corresponding to the 98.5 percentile of Student's t-distribution, and the default FieldTrip neighborhood definition and connectivity.

Transfer entropy

Transfer entropy (TE), a directional measure of information transfer sensitive to linear and non-linear coupling (Vicente et al., 2011), was used to determine the direction of interaction between the gastric slow-wave and the 10–11 Hz alpha power, at each of the source time series belonging to a significant PAC cluster. We computed TE between the alpha source power time series from the source-localized clusters and the EGG time series filtered $\pm$ 0.02 Hz of the EGG peak frequency using TRENTOOL (Lindner et al., 2011). The 12 min resting state data were segmented into 60 s segments for TE analysis, sampled at 20 Hz. The embedding delay and embedding dimension were estimated for each participant via Ragwitz' criterion (Ragwitz and Kantz, 2002), with the maximal value of each measure taken across participants as the optimal parameters. The time series were embedded using these parameters, and the TE value was computed at each voxel between the alpha source power and EGG signal. Each TE value was tested for statistical significance via the non-parametric statistical test provided by TRENTOOL. The number of TE interactions exceeding an arbitrary threshold of $p < 0.05$ uncorrected, either in the EGG→MEG, or MEG→EGG directions was tabulated for each participant, separately for the anterior and posterior significant PAC clusters. To test for an asymmetry of directional interactions, the number of pairs above this threshold was then compared between the two directions using a paired t-test, separately in the posterior and anterior clusters, with the results of both t-tests Bonferroni corrected for multiple comparisons.

Results

Existence of gastric-alpha coupling

For each participant, we determined gastric frequency at the EGG electrode showing the largest peak in the normal gastric range to take into account intersubject variability in stomach location and to maximize signal-to-noise ratio (Fig. 1a, mean EGG frequency 0.046 Hz $\pm$ 0.001 sem). We then computed coupling between the gastric phase and the amplitude of brain rhythms from 1 to 100 Hz using the modulation index (MI) (Tort et al., 2010). We compared the obtained MI values with estimated chance level (see Material and Methods) using a cluster-based procedure (Maris and Oostenveld, 2007) that intrinsically corrects at the group level for multiple comparisons across sensors, frequencies and time samples. Within the frequency range tested (1–100 Hz), significant gastric-brain coupling (Fig. 1b) occurred in the alpha range, at 10 and 11 Hz, in two bilateral parieto-occipital clusters with an extension over right fronto-temporal sensors (sum(t)=53.22, MonteCarlo p=0.0008, and sum(t)=52.57, Monte-Carlo $p=0.0008$, corrected for multiple comparisons). Summary statistics of gastric-brain coupling (Fig. 1c) show a distinct peak at 10 and 11 Hz, indicating that the effect is well localized to the alpha band. The topography of significant gastric-alpha coupling and alpha power overlap, but only partially (Fig. 1e).

Gastric-alpha coupling was highly specific to gastric frequency (Fig. 1d). We filtered the signal from the abdominal electrode with a center frequency slightly lower or higher than each participant's gastric frequency ($\pm$ 0.015 Hz, in steps of 0.005 Hz), and repeated the same PAC analysis. Clusters obtained with a slight offset from gastric frequency showed much smaller summary statistics, that decreased and became non-significant as the distance from the original EGG frequency increased. To determine if coupling between MEG power and EGG phase was sensitive to the individual

alpha peaking frequency of each participant (10.35 Hz ± 0.13, range 9.6–11.6 Hz), we recomputed the sensor-level statistics after aligning to each participant's alpha peak. We found that this did not modify the results (two significant bilateral clusters, sum(t)= 44.33, 34.16, MonteCarlo $p=0.0020$, 0.0064, with a similar topography). Individual phase-amplitude profiles of a subsample of participants (participants with largest, median and smallest MI) are plotted in Fig. 1f. Those profiles show that gastric-alpha coupling seem to involve both 1:1 and higher coupling modes since the PAC profile may show consistent deviations from a sinusoidal fit (Fig. 1f, left).

We then determined explained variance, i.e. the proportion of spontaneous alpha fluctuations explained by gastric phase. The rationale for determining explained variance is described in Fig. 2 (see also Material and methods). It relies on the comparison between the variance of the original alpha amplitude envelope and the variance of the alpha amplitude envelope sorted by gastric phase. We found that in the significant clusters gastric phase accounted for 8.0 ± 0.5% (range across participants: 4.4–12.1%) of the variance of alpha amplitude.

Control analyses

MI is in principle independent from power (Tort et al., 2010), but we nevertheless verified that gastric-alpha coupling was not driven by EGG nor alpha power. There was no significant correlation across participants between MI and 10–11 Hz power averaged across the significant clusters (Spearman $\rho=0.24$, $p=0.35$). EGG power did not correlate with MI either (Spearman $\rho=-0.34$, $p=0.178$). We also estimated the false positives that our statistical approach might generate. We tested whether any of the 1000 surrogate data sets created to estimate chance level could give rise to cluster statistics as large as those produced by original data. We did not find any surrogate data set where two clusters were as large as the two empirical clusters, thereby showing that the Monte-Carlo probability of obtaining the two empirical clusters by chance was smaller than 0.001. The probability of obtaining by chance a single cluster larger than one of the two original clusters was $p=0.0053$.

Gastric-alpha coupling occurs in parieto-occipital regions and right anterior insula

We then identified the cortical regions where significant gastric-alpha coupling takes place. We computed a time series of 10–11 Hz power per voxel per participant using a beamformer-based source localization, computed gastric-alpha coupling at each voxel using the MI and applied the same statistical approach as at the sensor level. Significant gastric-alpha coupling took place in two anatomical regions (Fig. 3 and Table 1). The posterior cluster (sum (t)=538.60, MonteCarlo $p=0.014$) comprised the parieto-occipital sulcus and calcarine fissure bilaterally. The anterior cluster (sum (t)=383.20, MonteCarlo $p=0.044$) was centered on the right anterior insula.

Directionality of interactions between stomach and brain

Lastly, we tested whether the stomach influenced the brain or vice-versa. Since the gastric rhythm is intrinsically generated in the stomach (Bozler, 1945; Suzuki et al., 1986; Sanders et al., 2014),

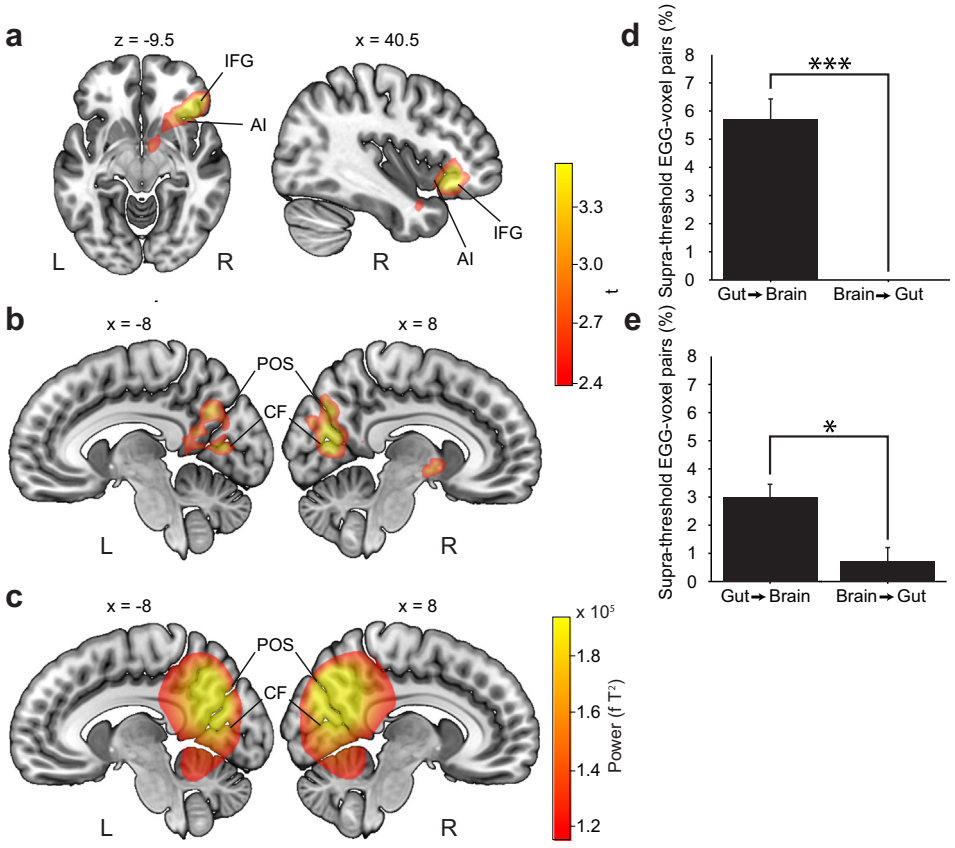


Fig. 3. Localization and directionality of significant gastric-alpha coupling. (a) Anterior cluster, centered on the right anterior insula (AI) and inferior frontal gyrus (IFG). (b) Posterior cluster, encompassing the parieto-occipital sulcus (POS) and calcarine fissure (CF) bilaterally. (c) Source-localization of alpha power, centered on the parieto-occipital sulcus and calcarine fissure bilaterally. (d) Causal interaction between stomach and brain is greater in the stomach-to-brain direction ($t(16)=7.98$, $p<10^{-5}$, corrected) in the right anterior insula. (e) Causal interaction between stomach and brain is greater in the stomach-to-brain direction ($t(16)=3.07$, $p=0.015$, corrected) in the posterior cluster.

Table 1
Anatomical description of the regions involved in gastric-alpha coupling, based on Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). Only areas with more than 1% of their volume involved are listed.

Cluster / AAL region	Peak t	t/mm^3	mm^3	Percent activation	MNI		
					X	*Y*	*Z*
Posterior							
Right Calcarine	5.21	3.59	3371	22.64	8	−68	16
Right Cuneus	5.11	3.45	1396	12.25	12	−68	20
Right Precuneus	4.91	3.26	1326	5.08	12	−67	21
Left Calcarine	4.56	3.23	1854	10.26	−8	−72	20
Left Cuneus	4.41	3.13	1158	9.49	−8	−72	23
Right Lingual Gyrus	4.28	3.28	502	2.73	7	−67	8
Right Inferior Occipital Gyrus	4.27	3.42	1528	19.31	44	−76	−4
Right Middle Occipital Gyrus	3.92	3.29	865	5.15	48	−76	0
Left Precuneus	3.81	3.08	1825	6.47	−4	−67	28
Left Superior Occipital Gyrus	3.79	3.02	177	1.62	−22	−64	24
Right Middle Temporal Gyrus	3.77	3.13	401	1.14	44	−73	−3
Anterior							
Inferior Frontal Gyrus, Orbital part	4.60	3.51	4423	32.39	36	28	−8
Right Insula	4.55	3.23	1527	10.78	36	28	−5
Inferior Frontal Gyrus, Triangular part	4.14	3.24	327	1.90	40	33	−3
Right Putamen	3.52	2.99	372	4.37	25	22	−8
Temporal Pole, Superior Temporal Gyrus	3.46	2.97	603	5.63	32	4	−24
Right Amygdala	3.16	2.78	276	13.91	36	2	−24
Right Olfactory	2.99	2.78	195	8.43	28	9	−20

we expected that the ascending direction, from stomach to brain, would predominate. We computed transfer entropy, a measure of directionality of information transfer, between the filtered EGG signal and amplitude envelope of the 10–11 Hz MEG signal, separately for the right anterior insula cluster and for the posterior parieto-occipital cluster. Information transfer was greatest from stomach to brain (Fig. 3) for both the parieto-occipital cluster ($t(16)=3.07$, $p=0.015$, Bonferroni corrected) and the anterior insula ($t(16)=7.98$, $p<10^{-5}$, Bonferroni corrected).

Discussion

We show here that the temporal structure of large-scale spontaneous brain dynamics is coupled with gastric signals. Gastric-brain coupling was revealed by a modulation of the amplitude of the alpha rhythm by gastric phase, in the parieto-occipital sulcus and calcarine fissure bilaterally and in the right anterior insula. These results show that the basic rule linking the phase of slow rhythms with the amplitude of higher frequency rhythms, so far observed only within the brain (Bragin et al., 1995; Canolty et al., 2006; Schroeder and Lakatos, 2009; Buzsaki, 2010), can be extended to interactions between brain and viscera. 8% of spontaneous alpha fluctuations were explained by gastric phase, and gastric-alpha coupling appears to be driven by ascending signals from stomach to brain.

We found that the largest component of spontaneous brain activity, the alpha rhythm, is locked to gastric phase. The alpha rhythm is known to exert an inhibitory influence on spike-firing rate (Haegens et al., 2011) and has a versatile impact on perception, attention and memory (Palva and Palva, 2007; Jensen and Mazaheri, 2010; Klimesch, 2012). Given the wide range of perceptual and cognitive correlates of alpha oscillations, the gastric rhythm might impose a slow temporal constraint over a range of processes, including basic stimulus detection that displays slow fluctuations (Monto et al. 2008) in the gastric frequency range.

Interestingly, the parieto-occipital regions where we find gastric-alpha coupling are not only associated with alpha rhythm generation (Salmelin and Hari, 1994), but they are also deactivated in response to experimentally-induced mechanical distension of the stomach, which leads to conscious and sometimes painful stomach sensations (van Oudenhove et al., 2009). In addition, electrical intraperitoneal stimulation elicits a response in the monkey visual cortex during sleep (Pigarev, 1994; Pigarev et al., 2006). The right anterior insula is also activated during gastric distension (Mayer et al., 2009) and is linked to gastric frequency changes during disgust (Harrison et al., 2010). Those experiments, that involve active stimulation of the stomach or emotional challenges, reveal the existence of anatomical circuits relaying visceral information to cortical structures, including occipito-parietal regions and right anterior insula. Our results show that during resting-state, in the absence of active gastric stimulation but in the presence of the gastric basal rhythm that is continuously generated, this circuitry is functional: the alpha rhythm in parieto-occipital regions and right anterior insula is coupled to the stomach.

fMRI studies have underlined the importance of bodily signals such as cardiac activity, respiration and blood pressure fluctuations, during the resting-state. However in this literature bodily signals are most often considered as artifacts injecting non-neural influences on the BOLD signal (Glover et al., 2000; Birn et al., 2006; Shmueli et al., 2007; Murphy et al., 2013). Direct measures of cerebral electrical activity, such as MEG or EEG, although not immune to physiological artifacts (Dirlich et al., 1997; Kern et al., 2013) can better reveal the coupling between bodily signals and neural activity. For instance, the brain transiently responds to heartbeats (Schandry and Montoya, 1996; Kern et al., 2013; Park et al., 2014; Babo-Rebelo et al., 2016, Babo-Rebelo et al. in press). The link between those transient responses and the temporal structure of large-scale spontaneous brain activity is not yet known, although there are reported interactions between heart timing and stimulus processing (Birren et al., 1963; Elliott and Graf, 1972; Gray et al., 2009; Garfinkel et al., 2014).

Here, we show that gastric activity is directly coupled to spontaneous neural activity. The directionality analysis we performed indicates that the transfer of information is predominantly in the stomach-to-brain direction, congruent with the fact that the gastric basal rhythm is intrinsically generated in the stomach (Sanders et al., 2006). We thus propose that the stomach could be considered as an external oscillator constraining spontaneous fluctuations of brain activity. This implies that the temporal structure of spontaneous brain activity depends not only on neuron and network properties (Buzsaki and Draguhn, 2004; Deco et al., 2009; Petersen and Sporns, 2015), but also on a slow oscillator in the stomach wall. So-called "intrinsic" brain dynamics might thus be better understood, modeled and reproduced (Hyafil et al., 2015; Ponce-Alvarez et al., 2015) by including visceral generators of rhythmic activity acting as external oscillators coupled to the brain.

Author contributions

C.R., M.B.R. and C.T.B. designed the experiment; C.R. and M.B.R. acquired the data; C.R., D.S and C.T.B. analyzed the data; C.R., M.B.R., and C.T.B. wrote the paper.

Acknowledgments

We thank Patricia Wollstadt, Michael Wibral, Jan-Mathijs Schoffelen, Robert Oostenveld, and Adriano Tort for assistance with data analysis, Christophe Gitton and Antoine Ducorps for assistance with data acquisition and preprocessing. This work has received funding from Agence Nationale de la Recherche ANR-BLAN-BSH2-0002-01 and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement no 670325) to CTB, as well as from Agence Nationale de la Recherche ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

References

Babo-Rebelo, M., Richter, C.G., Tallon-Baudry, C., 2016. Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. J. Neurosci. 36, 7829–7840.
Babo-Rebelo M., Wolpert N., Adam C., Hasboun D., Tallon-Baudry C., 2016. Is the cardiac monitoring function related to the self in both the default-network and right anterior insula? Phil. Trans. Roy. Soc. B. (in press).
Bahramisharif, A., van Gerven, M.A., Aarnoutse, E.J., Mercier, M.R., Schwartz, T.H., Foxe, J.J., Ramsey, N.F., Jensen, O., 2013. Propagating neocortical gamma bursts are coordinated by traveling alpha waves. J. Neurosci. 33, 18849–18854.
Birren, J.E., Phillips, S.L., Cardon, P.V., 1963. Reaction time as a function of cardiac cycle in young adults. Science 140, 195–196.
Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536–1548.
Bozler, E., 1945. The action potentials of the stomach. Am. J. Physiol. 144, 693–700.
Bragin, A., Jando, G., Nadasdy, Z., Hetke, J., Wise, K., Buzsaki, G., 1995. Gamma (40–100 Hz) oscillation in the Hippocampus of the behaving Rat. J. Neurosci. 15, 47–60.
Buzsaki, G., 2010. Neural syntax: cell assemblies, synapsembles, and readers. Neuron 68, 362–385.
Buzsaki, G., Draguhn, A., 2004. Neuronal oscillations in cortical networks. Science 304, 1926–1929.
Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger, M.S., Barbaro, N.M., Knight, R.T., 2006. High gamma power is phase-locked to theta oscillations in human neocortex. Science 313, 1626–1628.
Chen, J.D., Zou, X., Lin, X., Ouyang, S., Liang, J., 1999. Detection of gastric slow wave propagation from the cutaneous electrogastrogram. Am J Physiol 277, 424–430.
Critchley, H.D., Harrison, N.A., 2013. Visceral influences on brain and behavior. Neuron 77, 624–638.
Deco, G., Jirsa, V., McIntosh, A.R., Sporns, O., Kotter, R., 2009. Key role of coupling, delay, and noise in resting brain fluctuations. Proc. Natl. Acad. Sci. USA 106, 10302–10307.
Dirlich, G., Vogl, L., Plaschke, M., Strian, F., 1997. Cardiac field effects on the EEG. Electroencephalogr. Clin. Neurophysiol. 102, 307–315.
Elliott, R., Graf, V., 1972. Visual sensitivity as a function of phase of cardiac cycle. Psychophysiology 9, 357–361.
Florin, E., Baillet, S., 2015. The brain's resting-state activity is shaped by synchronized cross-frequency coupling of neural oscillations. Neuroimage 111, 26–35.
Furness, J.B., Rivera, L.R., Cho, H.J., Bravo, D.M., Callaghan, B., 2013. The gut as a sensory organ. Nat. Rev. Gastroenterol. Hepatol. 10, 729–740.
Garfinkel, S.N., Minatti, L., Gray, M.A., Seth, A.K., Dolan, R.J., Critchley, H.D., 2014. Fear from the heart: sensitivity to fear stimuli depends on individual heartbeats. J. Neurosci. 34, 6573–6582.
Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn. Reson. Med. 44, 162–167.
Gray, M.A., Rylander, K., Harrison, N.A., Wallin, B.G., Critchley, H.D., 2009. Following one's heart: cardiac rhythms gate central initiation of sympathetic reflexes. J. Neurosci. 29, 1817–1825.
Haegens, S., Nacher, V., Luna, R., Romo, R., Jensen, O., 2011. alpha-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. Proc. Natl. Acad. Sci. USA 108, 19377–19382.
Harrison, N.A., Gray, M.A., Gianaros, P.J., Critchley, H.D., 2010. The embodiment of emtional feelings in the brain. J. Neurosci. 30, 12878–12884.
Hyafil, A., Giraud, A.L., Fontolan, L., Gutkin, B., 2015. Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. Trends Neurosci. 38, 725–740.
Ito, S., 2002. Visceral region in the rat primary somatosensory cortex identified by vagal evoked potential. J. Comp. Neurol. 444, 10–24.
Jensen, O., Mazaheri, A., 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front. Hum. Neurosci. 4, 186.
Kern, M., Aertsen, A., Schulze-Bonhage, A., Ball, T., 2013. Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. Neuroimage 81C, 178–190.
Klimesch, W., 2012. alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn. Sci. 16, 606–617.
Koch, K.L., Stern, R.M., 2004. Handbook of Electrogastrography. Oxford University Press, Oxford.
Lindner, M., Vicente, R., Priesemann, V., Wibral, M., 2011. TRENTOOL: a Matlab open source toolbox to analyse information flow in time series data with transfer entropy. BMC Neurosci. 12, 119.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190.
Mayer, E.A., 2011. Gut feelings: the emerging biology of gut-brain communication. Nat. Rev. Neurosci. 12, 453–466.
Mayer, E.A., Aziz, Q., Coen, S., Kern, M., Labus, J.S., Lane, R., Kuo, B., Naliboff, B., Tracey, I., 2009. Brain imaging approaches to the study of functional GI disorders: a Rome working team report. Neurogastroenterol. Motil 21, 579–596.
Monto, S., Palva, S., Voipio, J., Palva, J.M., 2008. Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. J. Neurosci. 28, 8268–8272.
Murphy, K., Birn, R.M., Bandettini, P.A., 2013. Resting-state fMRI confounds and cleanup. Neuroimage 80, 349–359.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869.
Osipova, D., Hermes, D., Jensen, O., 2008. Gamma power is phase-locked to posterior alpha activity. PLoS One 3, e3990.
Palva, J.M., Palva, S., Kaila, K., 2005. Phase synchrony among neuronal oscillations in the human cortex. J. Neurosci. 25, 3962–3972.
Palva, S., Palva, J.M., 2007. New vistas for alpha-frequency band oscillations. Trends Neurosci. 30, 150–158.
Park, H.D., Correia, S., Ducorps, A., Tallon-Baudry, C., 2014. Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat. Neurosci. 17, 612–618.
Petersen, C.C., Sporns, O., 2015. Brain networks and cognitive architectures. Neuron 88, 207–219.
Pigarev, I., Almirall, H., Pigareva, M.L., Bautista, V., Sanchez-Bahillo, A., Barcia, C., Herrero, M.T., 2006. Visceral signals reach visual cortex during slow wave sleep: study in monkeys. Acta Neurobiol. Exp. 66, 69–73.
Pigarev, I.N., 1994. Neurons of visual cortex respond to visceral stimulation during slow wave sleep. Neuroscience 62, 1237–1243.
Ponce-Alvarez, A., Deco, G., Hagmann, P., Romani, G.L., Mantini, D., Corbetta, M., 2015. Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. PLoS Comput. Biol. 11, e1004100.
Powley, T.L., Phillips, R.J., 2011. Vagal intramuscular array afferents form complexes with interstitial cells of Cajal in gastrointestinal smooth muscle: analogues of muscle spindle organs? Neuroscience 186, 188–200.
Ragwitz, M., Kantz, H., 2002. Markov models from data by simple nonlinear time series predictors in delay embedding spaces. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 65, 056201.
Riezzo, G., Russo, F., Indrio, F., 2013. Electrogastrography in adults and children: the strength, pitfalls, and clinical significance of the cutaneous recording of the gastric electrical activity. Biomed. Res. Int. 2013, 282757.
Roux, F., Wibral, M., Singer, W., Aru, J., Uhlhaas, P.J., 2013. The phase of thalamic alpha activity modulates cortical gamma-band activity: evidence from resting-state MEG recordings. J. Neurosci. 33, 17827–17835.
Salmelin, R., Hari, R., 1994. Characterization of spontaneous MEG rhythms in healthy adults. Electroencephalogr. Clin. Neurophysiol. 91, 237–248.
Sanders, K.M., Koh, S.D., Ward, S.M., 2006. Interstitial cells of Cajal as pacemakers in the gastrointestinal tract. Annu. Rev. Physiol. 68, 307–343.
Sanders, K.M., Ward, S.M., Koh, S.D., 2014. Interstitial cells: regulators of smooth muscle function. Physiol. Rev. 94, 859–907.
Schandry, R., Montoya, P., 1996. Event-related brain potentials and the processing of cardiac activity. Biol. Psychol. 42, 75–85.
Schroeder, C.E., Lakatos, P., 2009. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 32, 9–18.
Shmueli, K., van Gelderen, P., de Zwart, J.A., Horovitz, S.G., Fukunaga, M., Jansma, J. M., Duyn, J.H., 2007. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. Neuroimage 38, 306–320.
Suzuki, N., Prosser, C.L., Dahms, V., 1986. Boundary cells between longitudinal and circular layers - essential for electrical slow waves in cat intestine. Am. J. Physiol. 250, G287–G294.
Tort, A.B., Komorowski, R., Eichenbaum, H., Kopell, N., 2010. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. J. Neurophysiol. 104, 1195–1210.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.
van Oudenhove, L., Vandenberghe, J., Dupont, P., Geeraerts, B., Vos, R., Bormans, G., van Laere, K., Fischler, B., Demyttenaere, K., Janssens, J., Tack, J., 2009. Cortical deactivations during gastric fundus distension in health: visceral pain-specific response or attenuation of 'default mode' brain function? A H2 15O-PET study. Neurogastroenterol. Motil. 21, 259–271.
Vicente, R., Wibral, M., Lindner, M., Pipa, G., 2011. Transfer entropy–a model-free measure of effective connectivity for the neurosciences. J. Comput. Neurosci. 30, 45–67.
Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., Pennartz, C.M., 2010. The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. Neuroimage 51, 112–122.
Weaver, K.E., Wander, J.D., Ko, A.L., Casimo, K., Grabowski, T.J., Ojemann, J.G., Darvas, F., 2016. Directional patterns of cross frequency phase and amplitude coupling within the resting state mimic patterns of fMRI functional connectivity. Neuroimage 128, 238–251.

B. Article II: The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision

Cortex, Special Issue: Unconscious guidance of attention

The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision.

Catherine Tallon-Baudry[1], Florence Campana[1,2], Hyeong-Dong Park[1,3], Mariana Babo-Rebelo[1]

1- Cognitive Neuroscience Laboratory, Inserm u960, Ecole Normale Supérieure, Paris, France

2- The Dynamic Perception Lab, Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, MD, USA

3- Laboratory of Cognitive Neuroscience, Center for Neuroprosthetics and Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland

Corresponding author:

Catherine Tallon-Baudry, LNC, 29 rue d'Ulm 75005 Paris, France

catherine.tallon-baudry@ens.fr

Keywords: vision, perception, attention, consciousness, heartbeat evoked response, default network, heart, self

Abstract

Why should a scientist whose aim is to unravel the neural mechanisms of perception consider brain-body interactions seriously? Brain-body interactions have traditionally been associated with emotion, effort, or stress, but not with the "cold" processes of perception and attention. Here, we review recent experimental evidence suggesting a different picture: the neural monitoring of bodily state, and in particular the neural monitoring of the heart, affects visual perception. The impact of spontaneous fluctuations of neural responses to heartbeats on visual detection is as large as the impact of explicit manipulations of spatial attention in perceptual tasks. However, we propose that the neural monitoring of visceral inputs plays a specific role in conscious perception, distinct from the role of attention. The neural monitoring of organs such as the heart or the gut would generate a subject-centered reference frame, from which the first-person perspective inherent to conscious perception can develop. In this view, conscious perception results from the integration of visual content on the one hand, and of the subject-centered reference frame on the other hand.

Does it matter that the brain is embedded in a body to understand vision? Leaving the oculo-motor system aside, it is usually held that basic mechanisms of perception are independent from bodily influences except under special conditions of stress, arousal or emotion. We review here recent evidence suggesting a quite different picture, and explain how the neural monitoring of bodily signals could fill an important gap in our understanding of conscious vision. In the first part of the article, we argue that to understand how a conscious percept is formed, it is not sufficient to consider perceptual mechanisms and higher cognitive functions such as attention and memory. A simple but core component of conscious perception, first-person perspective, has to be accounted for. In the second part, we present the hypothesis that first-person perspective derives from a subject-centered reference frame. This egocentric reference frame would be created by the neural monitoring of visceral organs. We review the recent experimental evidence supporting the hypothesis that to account for the statement "I have seen the stimulus", a neural model should not only describe mechanisms related to perceptual detection and decision making, but also propose a mechanism to explain where the "I" is coming from.

I. Perceptual consciousness: neither attention nor high-level cognition

Consciousness has long been conceived as an overarching cognitive function associated with high-level, finely tuned behavior. In an influential pioneering model (Baars, 1997), Baars defined consciousness as the spotlight of attention shining on the stage of working memory. In the following 20 years, a large number of experimental studies investigated the links between perceptual consciousness and attention and, to a lesser extent, working memory. In the section below, we review the arguments showing that those two high-level cognitive functions cannot explain perceptual consciousness. We argue that it is time to concentrate on another, core aspect of consciousness: first-person perspective or subjectivity.

Attention is distinct from consciousness

The idea that attention drives consciousness is appealing (Dennett, 1991; Dehaene and Naccache, 2001) and fits with numerous behavioral observations. For instance, attention

facilitates detection (Solomon, 2004) and enhances perceived contrast (Carrasco et al., 2004). Conversely, in the absence of attention, salient stimuli may not be reported, as in inattentional blindness (Mack and Rock, 1998), change blindness (Simons and Levin, 1997), or during the attentional blink (Shapiro et al., 1997). It thus seems that both attention and consciousness correspond to "perceiving better".

However, the fact that attention facilitates the report "I have seen the stimulus" that is the hallmark of visual consciousness does not imply that attention and consciousness are the same. Rather, attention and consciousness correspond to distinct neural mechanisms, that can both independently contribute to the final decision of reporting the presence or absence of the stimulus (Tallon-Baudry, 2012). This view is anchored in a growing number of experimental findings teasing apart the neural correlates of attention and consciousness and their behavioral consequences. In the past 10 years, the neural correlates of attention and consciousness could be repeatedly dissociated, either partly (Koivisto et al., 2006; Watanabe et al., 2011; Webb et al., 2017) or fully (Schurger et al., 2008; Wyart and Tallon-Baudry, 2008; Wyart et al., 2012). In parallel, a growing number of behavioral experiments showed that attention can be triggered by unconscious cues or affect unconsciously processed targets (see e.g., (Kentridge et al., 1999; Kentridge et al., 2004; Norman et al., 2013)).

While the idea that attention and consciousness should not be conflated gained strength, it is still sometimes argued that attention is a gate for consciousness. If this were the case, attention should *always* facilitate consciousness, which is contradicted by three lines of findings. Firstly, the neural correlates of consciousness do not necessarily depend on attention (Koivisto et al., 2006; Wyart and Tallon-Baudry, 2008; Wyart et al., 2012). Secondly, attention and consciousness can have opposite behavioral consequences (van Boxtel et al., 2010). Lastly, the conscious or unconscious status of the stimulus can determine the type of attention deployed, reversing the link of causality between attention and consciousness (Hsu et al., 2011).

Other cognitive functions thought to be tightly associated with consciousness have seen their status revised. Neural markers of semantic information processing can be measured in response to unseen words (Luck et al., 1996). Unperceived stimuli can be maintained in short-term memory (Soto et al., 2011; Sergent et al., 2013; King et al., 2016). Unconscious errors are detected by the anterior cingulate cortex (Hester et al., 2005). The

frontal activations that were once thought to be markers of consciousness can contribute to unconscious and involuntary control (Lau and Passingham, 2007; Sumner et al., 2007; van Gaal et al., 2008) and are associated with behavioral report, rather than with conscious perception *per se* (Frässle et al., 2014).

Attention, memory and control can thus operate on unconscious stimuli, and might be influenced by consciousness rather than driving it. It follows that cognitive functions such as attention, memory and control, cannot explain conscious perception. It is thus time to reconsider the nature of perceptual consciousness.

First-person perspective and subjectivity as core components of consciousness

Experimentally, the hallmark of conscious vision is the report "I have seen the stimulus": it implies the existence of a subject, with his or her own first-person perspective, who can say "I". The point we want to make in this article is that to account for the statement "I have seen the stimulus", a neural model should not only describe mechanisms related to perceptual detection and decision making, but also propose a mechanism to explain where the "I" is coming from (Park and Tallon-Baudry, 2014). The combination of first-person perspective with visual content would give rise to subjective experience (Figure 1).

The famous picture associated with #TheDress provides a good example of what we mean by subjectivity in conscious vision. Depending on individuals, the dress presented in this picture can be perceived as blue and black or white and gold. A likely explanation for inter-individual differences in the perception of this image is that different participants interpret differently the nature of ambient light in the picture, and hence perceive colors differently (Lafer-Sousa et al., 2015). However, these mechanistic explanations do not account for the frustration and rage expressed by social media users at discovering the subjectivity of perception. Indeed, as opposed to most well-known ambiguous images such as the duck-rabbit or face-vase illusions, the perceived color of #TheDress does not change over time: a blue-black perceiver cannot know "how it feels like" to perceive the dress as white and gold. This example underlines the importance of the first-person perspective of the experiencing subject to fully account for conscious visual perception.

Subjectivity has long been banished from scientific investigations as a notion that cannot be addressed scientifically, since it is by essence private. This inheritance from behaviorism should not hide the fact that subjectivity is an ingredient key to conscious perception. Indeed, conscious perception can only exist if the stimulus is experienced by a subject. This point has long been emphasized by philosophers (Nagel, 1974; Chalmers, 1995; Searle, 2000; Zahavi, 2003; Block, 2007) but subjectivity remains absent from mechanistic models of conscious vision and attention.

Subjective experience is not an illusion

The very existence of subjective experience has been denied (Dennett, 1991; O'Regan and Noe, 2001): subjective experience would be a post-hoc cognitive reconstruction rather than an immediate experience (Dehaene et al., 2006; Cohen and Dennett, 2011). Indeed, spontaneous subjective reports do not always survive scientific scrutiny. In change blindness studies for instance (Simons and Levin, 1997; Rensink, 2002), two images of the same visual scene differing by one item are presented in rapid succession, separated by a blank screen. Although the change can be massive, it often remains unnoticed. In other words, subjects have the feeling they see the entire visual scene – a rich subjective experience – but when probed they are unable to report accurately the details of the visual scene. We have previously (Campana and Tallon-Baudry, 2013) pointed out that while the experimental manipulation used in change blindness paradigm does indeed prevent the conscious perception of details, the subject may nevertheless truly perceive consciously the gist of a visual scene. In this view presented in the left part of Figure 1 and based on the influential reverse hierarchy theory (Hochstein and Ahissar, 2002), local details would first be processed unconsciously in early visual areas and rapidly combined, in a feed-forward and automatic manner, into a global scene in higher-order visual areas. The result of this first wave of computation is the gist of the scene and can be perceived consciously. The conscious perception of local details would require an additional and optional processing step proceeding from higher-order to lower-order areas.

The reverse hierarchy theory (Hochstein and Ahissar, 2002) was initially proposed to account for findings in perceptual learning but fits with experimental findings in the domain of perception and attention, such as the fact that attention proceeds from

higher-order to lower-order visual areas (Luck et al., 1997; Mehta et al., 2000; Buffalo et al., 2010). However, the crucial prediction that conscious percepts are preferentially formed at a global level remained to be validated. To test this prediction, we designed new stimuli that are truly hierarchical, as opposed to the classic Navon's letters (Navon, 1977; Kimchi, 1992). Stimuli were composed of local and global information that could be varied independently, but where global information existed only by virtue of local information (Campana et al., 2016). We verified three key predictions. Firstly, participants respond faster when instructed to respond on global features than when instructed to respond on local features, showing that global information is easier to access than local details. Secondly, global information is computed by the brain irrespective of task demands, in line with the hypothesis that global information is automatically computed during the fast feed-forward sweep. Lastly, spontaneous reports were dominated by global information, in line with the hypothesis that conscious percepts are preferentially formed at a global level.

Conscious percepts are thus formed preferentially at the global level, and the conscious identification of local details is optional and time consuming, as could be predicted by the reverse hierarchy theory (Hochstein and Ahissar, 2002; Campana and Tallon-Baudry, 2013). It follows that this model offers an alternative and parsimonious explanation of the experimental findings in the change blindness paradigm: participants truly experience the gist of the scene but are prevented to further analyze local details because of time pressure and masking effects. Subjective experience is thus not an illusion, its properties derive from the architecture of the visual system. However, the architecture of the visual system by itself does not account for subjective experience. How is subjective experience implemented?

II. Accounting for the "I" in the report "I have seen the stimulus"

A first step to account for subjective experience would be to tag some neural activities as being related to the subject of the experience, the "I". It could be argued that there is no need for a *specific* mechanism related to the "I", since *any* neural process taking place inside the brain could be labeled as belonging to the organism, as being "I-related". Earlier in this article, we have reviewed evidence that elaborated visual and cognitive processing can take place unconsciously, i.e. without subjective experience. Hence, neural

processes are not equipped by default with the "I-relatedness" necessary for conscious experience.

A subject-centered reference frame based on visceral inputs to account for first-person perspective

What type of signals could be good candidates for establishing subjectivity? Signals originating in the body and relayed up to the brain could be self-specifying, since they could provide the brain with a definition of the organism. Bodily signals have been proposed to play a role in the emergence of subjectivity (Gallagher, 2000; Zahavi, 2003; Craig, 2009; Damasio, 2010). Experimental studies on agency (David et al., 2008) or bodily awareness (Petkova et al., 2011; Blanke, 2012; Ferre et al., 2014) underline the role of sensory signals from the skin, limbs, joints or vestibular system. However, this type of bodily afferences cannot be sufficient. Locked-in patients, who are fully paralyzed and whose brain does not receive any feedback on bodily movement or action performance, are nevertheless conscious (Tononi and Koch, 2008).

The brain has other major sources of bodily information: the viscera, that include organs such as the heart and the gut, constitute another excellent but overlooked candidate. Both the gut and the heart are pacemakers, in the sense that they generate their own electrical activity. While the pacemaker activity of heart is well known, the discovery that the digestive tract is lined with a specific cell type that intrinsically and continuously generates a slow electrical rhythm is more recent (Kelly and Code, 1971; Furness, 2006; Sanders et al., 2006).

The potential role of this *ascending* information, from viscera to the neocortex, has been little explored so far (Critchley and Harrison, 2013; Park and Tallon-Baudry, 2014; Richter et al., 2017). Rather, most neuroimaging studies focused on *descending* commands from the central nervous system that regulate cardiac function (Wong et al., 2007; Thayer et al., 2012; Beissner et al., 2013), and control autonomic outputs such as skin conductance level (Nagai et al., 2004; Fan et al., 2012) or pupil diameter (Murphy et al., 2014).

We propose to view both the gut and heart as ticking clocks that constantly send intrinsically-generated ascending information up to the central nervous system. They could

thus provide a stable source of signals defining the organism as an entity at the neural level. The monitoring of those signals by the brain would thereby create an ego-centric, self-centered neural reference frame (Figure 1, right) from which first-person perspective can develop (Park and Tallon-Baudry, 2014).

Note that in this mechanistic framework, ascending signals from visceral organ do not have to be consciously perceived. The role of visceral afferents is here purely mechanistic: visceral organs feed the brain with signals that become self-specifying when reaching the brain, and lay the basis for a self-centered referential that is not experienced as such. Importantly, visceral signals do not have to necessarily indicate a change in bodily state to contribute to conscious perception, as opposed to their proposed role in influential theories on the self (Craig, 2002; Damasio, 2010) or on emotion, such as the James-Lange theory or the somatic marker hypothesis (Damasio, 1996).

Neural responses to heartbeats

Cognitive neuroscience is fortunately equipped with a powerful tool to study the neural monitoring of the heart: heartbeat-evoked responses (Schandry and Montoya, 1996). Heartbeat evoked responses are obtained by averaging electrophysiological data time-locked to heartbeats (Schandry et al., 1986). They can thus be considered as equivalent to classical evoked responses obtained by time-locking data to the presentation of a visual or auditory stimulus, but in this instance, the stimulus is internal. It is also important to bear in mind the presence of an associated cardiac artefact, because sensors on the head pick up not only the neural response to heartbeats, but also the electro-cardiogram (Dirlich et al., 1997).

Heartbeat evoked responses share a number of properties similar to classical sensory responses. Heartbeat evoked responses are modulated by attention, e.g. when participants have to count the occurrence of their own heartbeats, and the amplitude of the heartbeat evoked response relates to accuracy at the heartbeat counting task (Schandry et al., 1986; Montoya et al., 1993; Pollatos and Schandry, 2004; Canales-Johnson et al., 2015). The amplitude of the heartbeat evoked response depends on participant's state as measured by alpha power (Luft and Bhattacharya, 2015) or by sleep stages (Lechinger et al., 2015).

Heartbeat evoked responses are also modulated by emotion (Fukushima et al., 2011; Couto et al., 2015).

A possible origin of neural responses to heartbeats is the neural discharge of the mechano-receptors in the heart wall and aortic arch (Shepherd, 1985; Armour and Ardell, 2004). Those mechanoreceptors discharge at each cardiac cycle in response to the mechanical distortions of the cardiac and aortic walls. This information is relayed, through spinal and vagal pathways, to the nucleus tractus solitarius, the parabrachial nucleus and to the thalamus. Those nuclei in turn target a number of structures: the amygdala, the cerebellum, hypothalamus, locus coeruleus and nucleus accumbens, but also cortical structures such as the primary and secondary somatosensory cortices, the insula or the ventral anterior cingulate / ventro-medial prefrontal cortex (vACC-vmPFC) (Vogt et al., 1987; Pritchard et al., 2000; Henry, 2002; Van der Werf et al., 2002; Critchley and Harrison, 2013). A direct spinal projection could recently be traced up not only to insular and secondary somatosensory cortex, but also to cingulate motor areas in monkeys (Dum et al., 2009). It is worth mentioning that the mechanisms and pathways underlying neural responses to heartbeats are so far poorly characterized. The mechano-receptor hypothesis is compatible with known physiology and anatomy, but it has not been investigated directly by recording, for instance, neural discharges in response to heartbeats in nucleus tractus solitarius and cortical target sites. Recent findings suggest that other mechanisms might also play a role. For instance, somatosensory signals from the skin in the heart region might also contribute (Khalsa et al., 2009). It has also recently been discovered that local changes in blood pressure provoke changes in spontaneous neural firing in rodent slices (Kim et al., 2016), suggesting that vascular events can directly affect neural activity, at least *in vitro*.

While the insula has often been presented as the primary visceral area, anatomy suggests a much more distributed pattern, as described in the preceding paragraph. Functional results confirm a distribution of neural responses to heartbeats in several regions predicted by anatomical pathways. In human intra-cranial recordings, neural responses to heartbeats have been observed in the primary somato-sensory cortex (Kern et al., 2013), the vACC-vmPFC and the insula (Babo-Rebelo et al., 2016b). Source reconstruction of magneto-encephalographic (MEG) and electro-encephalographic (EEG) data points to neural responses to heartbeats in vACC-vmPFC (Park et al., 2014; Babo-Rebelo et al., 2016a) and

mid-cingulate motor cortex (Park et al., 2016). Functional responses have also been observed in other regions not directly predicted by anatomical pathways, such as the right angular gyrus (Park et al., 2014) and posterior cingulate cortex (Babo-Rebelo et al., 2016a; Babo-Rebelo et al., 2016b).

Heartbeat-evoked response and perception of gratings at threshold

Neural responses to heartbeats before stimulus onset predict whether a faint stimulus at detection threshold will be perceived or missed. We (Park et al., 2014) presented participants with gratings at threshold for detection (Figure 2A). At each trial, when participants fixated properly, the fixation mark turned red to indicate the beginning of the trial. After a variable delay, a grating could appear or not. Participants were simply required to indicate at the end of the trial whether a stimulus had been presented or not.

A classical approach to such an experiment would be to focus on perceptual processing and decision making. Here, we adopted a different perspective. Our hypothesis was that part of the fluctuations between the report 'I have seen the stimulus' (hits) and the report 'I have not seen anything' (misses) is related to fluctuations of the 'I' and would be indexed by fluctuations in neural responses to heartbeats. We thus analyzed heartbeat evoked responses, measured with magneto-encephalography, before stimulus onset, during the warning interval (Figure 2B). Heartbeat evoked responses were larger in hits than in misses or correct rejections. Differential heartbeat evoked responses originated from the ventral anterior cingulate cortex / ventro-medial prefrontal cortex region (vACC-vmPFC) as well as from the right inferior parietal lobule (rIPL). The difference between heartbeat evoked responses in hits and misses remained below statistical threshold in the right insula. The difference in neural responses to heartbeats occurred at a moment when there was no difference between hits and misses in none of the cardio-respiratory parameters we measured (electrocardiogram, heart rate, blood pressure, respiration rate and phase).

The heartbeat evoked response co-varies with perceptual sensitivity, not decision criterion nor arousal

The amplitude of heartbeat evoked responses before stimulus onset accounts for a modulation of the hit rate. But what does this modulation reflect? We first checked that the results did not reflect a global, non-specific difference in arousal state between hits and misses. There was no evidence that arousal differed between hits and misses before stimulus onset: neither alpha power, nor pupil diameter, nor any of the measured cardio-respiratory parameters revealed any difference. In addition, the visual response to the warning stimulus was identical in hits and misses, suggesting that the larger responses to heartbeats in hits were not the result of a general, non-specific increase in cortical reactivity.

We then tested whether neural responses to heartbeats co-varied with sensitivity or criterion, and found clear-cut evidence that neural responses to heartbeats co-vary with perceptual sensitivity, not with decision criterion (Figure 2C). In addition, the size of the effects of neural responses to heartbeats on sensitivity and hit rate were similar to the effects of spatial attention that we observed in previous experiments. We found that the amplitude of the neural response to heartbeats accounts for 5 to 10 points of hit rate and for an 8% increase in sensitivity. Using similar gratings at threshold and manipulating spatial attention, we found in previous experiments that endogenous spatial attention modulates hit rate by 9 points (Wyart and Tallon-Baudry, 2008), and that exogenous spatial attention modulates sensitivity by 6% (Sergent et al., 2013). To summarize, neural activity obtained in response to heartbeats is used as sensory evidence in the final decision, and has as much influence on perceptual behavior as spatial attention would. But is it attention, or a neural marker of the "I"?

Neural responses to heartbeats index self-relatedness

Let us consider an attentional interpretation. It is known that when participants pay attention to their heartbeats, the amplitude of the heartbeat-evoked response increases (Schandry and Montoya, 1996). It seems unlikely that participants were counting or explicitly paying attention to their heartbeats while attempting at detecting a grating at threshold. Besides, interoceptive attention modulates activity in the insula (Critchley et al., 2004), whereas we found the largest differential responses to heartbeats in vACC-vmPFC and rIPL. However, it might be that participants' attention sometimes wandered away from the task and the screen, and turned inwards, to internal, task-unrelated thoughts. Such an "attention

inward" situation would lead to both larger responses to heartbeats and a greater probability of missing the stimulus displayed on screen. This interpretation does not fit with the observed data: larger responses to heartbeats were associated with an increase in hit rate, not with an increase in miss rate as predicted by the "attention inward" interpretation.

If neural responses to heartbeats are not related to an attentional effect, how can they behave as sensory evidence? Neural responses to heartbeats co-vary with visual sensitivity but are neither directly related to visual processing, since it occurs in response to heartbeats outside the visual system, nor directly related to the attentional modulation of visual processing. To interpret this intriguing finding, it is useful to explicitly formulate the statement that corresponds to hits and misses. In response to the same physical stimulus, participants report "I have seen the grating" in hits, and "I have not seen anything in misses". The classical approach to determine the neural mechanisms leading to such a statement focuses on perceptual and decisional processes. We suggest here that neural responses to heartbeats might have something to do with the "I" part of the sentence, with the fact that this statement comes from a subject having an experience: saying "I have seen the stimulus" implies the existence of the first-person perspective of the experiencing subject (Park and Tallon-Baudry, 2014).

This interpretation is strengthened by a series of recent experiments pointing toward a direct link between cardiac inputs and the self. We ran an interrupted thought experiment, where participants could let their mind wander freely but were interrupted from time to time and asked to rate the self-relevance of the current thought. In two separate experiments using either MEG in healthy participants (Babo-Rebelo et al., 2016a) or intracranial EEG in epileptic patients (Babo-Rebelo et al., 2016b), we found that neural responses to heartbeats indexed self-relevance (Figure 2 D). Self-relevance was defined as thinking *about oneself*, such as in "I am thirsty", or as being *the subject* experiencing or acting in the thought, such as "I will go to the supermarket this evening". Neural responses to heartbeats in vmPFC varied depending on whether the participant was thinking about himself/herself, or about an external object or event. This effect could be reproduced in intracranial recordings in vmPFC, with a significant correlation between the content of a single thought and the amplitude of neural responses to heartbeats in vmPFC during that thought (Figure 2E). Neural responses in the posterior cingulate cortex co-varied with more

experiential or agentive aspects of the self. This aspect of the self is pre-reflective, in the sense that one usually thinks about going to the supermarket without reflecting on oneself, but rather concentrating on the list of groceries. This form of pre-reflective self is always present in conscious mental life, but can be more (as in "I will call the travel agency") or less pronounced (as in "It's raining"). Note that in this interrupted thought paradigm, participants also rated their thoughts according to their emotional content. Neural responses to heartbeats did not vary with the emotional rating.

Neural responses to heartbeats do thus vary with self-relevance. In addition, in both the interrupted thought experiment and in the perception at threshold experiment, neural responses to heartbeats took place in the midline nodes of the default-network, that have been repeatedly associated with the self in fMRI (Qin and Northoff, 2011). Note that in those experiments (Park et al., 2014; Babo-Rebelo et al., 2016a; Babo-Rebelo et al., 2016b), the involvement of the insula was limited, remaining below statistical threshold. It might be that the insula is more involved in explicit interoception tasks, such as when participants are instructed to detect their heartbeats, and less so in the more automatic, unconscious monitoring process we targeted.

Another line of evidence for a link between neural responses to heartbeats and self-relatedness comes from studies on the bodily self. The experience of body ownership can be modulated by manipulating the synchrony between visual and tactile inputs. When those external stimuli are synchronized with the timing of heartbeats, illusions are enhanced (Aspell et al., 2013; Suzuki et al., 2013; Sel et al., in press), and neural responses to heartbeats co-vary with illusion strength (Park et al., 2016; Sel et al., in press). The modulation of heartbeat-evoked responses related to the bodily self takes place in midline motor and premotor regions (cingulate motor areas, supplementary motor area) (Park et al., 2016).

III. Conclusion, limitations and future directions

We have shown on the one hand that the sensory representations most likely to give rise to a conscious percept have a global, integrated content, corresponding to high-level visual areas. On the other hand we have provided evidence that neural responses to

heartbeats predict conscious perception by signaling a simple form of self, in particular in the default network. We propose that subjective experience results from the integration of visual content with the self-relatedness provided by neural responses to heartbeats (Figure 1). Note that other signals might contribute to this egocentric reference frame, for instance the stomach that intrinsically generates an electrical rhythm that impacts brain dynamics (Richter et al., 2017), as well as proprioceptive and vestibular inputs (Blanke, 2012). Besides, while we focused here on perception, our proposal can in principle extend to any cognitive process that includes a subjective aspect, for instance subjective value in value-based decision-making, or emotional appraisal.

From correlation to causation

The evidence presented in this article is correlational: neural responses to heartbeats before stimulus onset correlate with visual sensitivity, neural responses to heartbeats correlate with self-relevance. Moving from correlation to causation is an important but notoriously difficult step. Two approaches can be considered: altering viscera-to-brain communication, or establishing a model that generates new predictions that can be tested experimentally.

In our experiments, cardiac parameters did not vary; rather, neural responses to heartbeats varied in the absence of measured changes in cardio-respiratory parameters. While fluctuations in some cardiac parameters may not have been adequately measured, this suggests that neural variability, rather than cardiac variability, is crucial. To probe whether the neural monitoring of visceral signals plays causal role, the critical targets are thus the viscera-to-brain pathways and the central monitoring of visceral signals. Ascending pathways, from viscera to brain, can follow two routes, spinal and vagal, that are unlikely to be both severed in patients. In addition, both spinal and vagal pathways convey information in both directions, from viscera to brain and from brain to viscera. It follows that neither vagus nerve stimulation nor vagotomy specifically target ascending pathways. Besides, both interventions leave the spinal pathway intact. Still, it is worth mentioning that bariatric surgery, that usually implies vagotomy, increases the risk of self-harm behavior (Tindle et al., 2010; Bhatti et al., 2016), which may point to an underlying disturbance of the self.

To establish a model of conscious vision that generates new predictions that can be tested experimentally, one needs to identify when and how sensory evidence in the visual system is combined with the self-related information carried by neural responses to heartbeats in the default network. This question taps onto the general issue of large-scale information integration in the brain, that is far from being solved. Still, if a specific integration mechanism, be it convergence in a given area or oscillatory synchrony at a given frequency between two areas, were identified, this mechanism could be selectively disrupted using transcranial magnetic stimulation for instance.

Links with other experimental findings

We have reviewed evidence relating neural responses to heartbeats to the signaling of self-relatedness. Other types of cardiac-related effects have been described, called cardiac cycle effects or cardiac synchrony effects. Cardiac activity occurs in cycles of contraction and relaxation of the atria and ventricles, as reflected by the peaks of the electrocardiogram (ECG). Between the R and T peaks of the ECG, blood pressure is maximal (systole), as opposed to later in the cardiac cycle where blood pressure decreases (diastole). Supra-threshold stimuli presented during systole vs. diastole (for review, (Park and Tallon-Baudry, 2014); see also (Salomon et al., 2016)) are detected at different speeds in simple reaction time tasks and generate sensory evoked responses of different sizes. Oculo-motor behavior also depends on the cardiac cycle, with an excess of micro-saccades and fixational drifts at short latencies after the R peak, during systole (Ohl et al., 2016). However, peri-threshold stimuli tell a different story. The detection of neither visual (Elliott and Graf, 1972) nor auditory (Delfini and Campos, 1972; Velden and Juris, 1975) stimuli at threshold depend on the timing of the stimulus with respect to the cardiac cycle. In line with those findings, in our experiment on vision at threshold (Park et al., 2014), the perceived or unperceived fate of the stimulus did not depend on its position in the cardiac cycle. It remains to be determined whether the self-relatedness expressed by neural responses to heartbeats and the cardiac-cycle effects can be reconciled in the same framework.

Conclusion

We propose here that conscious visual experience results from the integration of visual content with an egocentric, self-related reference frame based on the neural monitoring of visceral organs (Figure 1). This proposal accounts for both the "I-related" and vision-related aspects of the report "I have seen the stimulus" that is the hallmark of visual consciousness. While recent experimental evidence support our proposal, the mechanism integrating visual information and neural responses to heartbeats, that are encoded in distinct brain areas, remains to determined. In this framework, attention is not the selection process that brings some items to the conscious mind, but rather a prioritization computational process that can operate on, or be triggered by, either consciously perceived or unconsciously processed stimuli. It follows from our proposal that even for a "cold" process such as vision at threshold, the fact that the brain is embedded into a body matters.

Acknowledgements.

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 670325) to CTB, as well as from ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*.

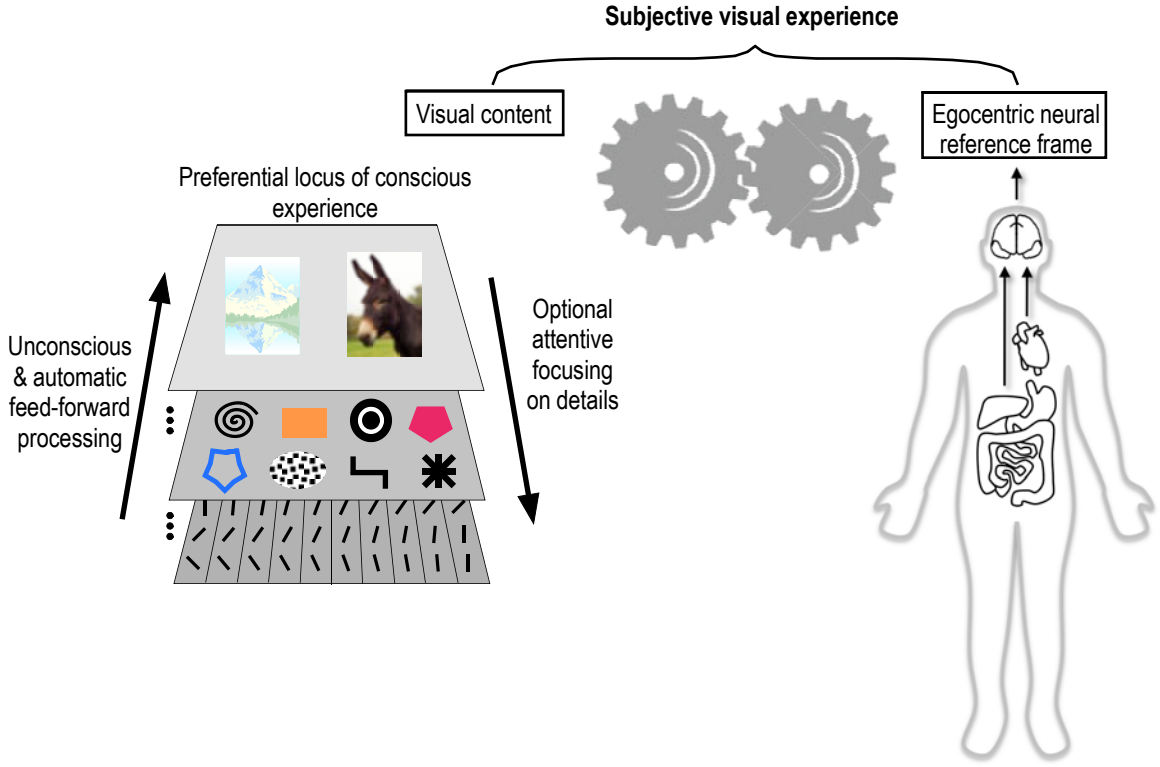
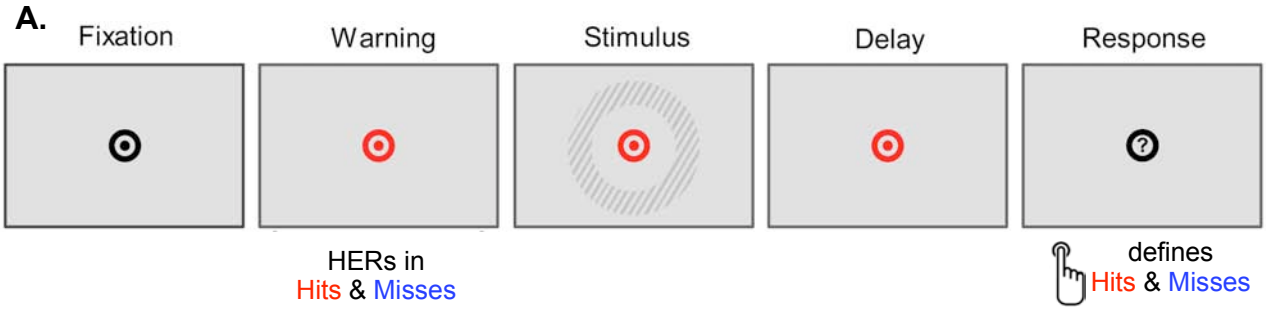
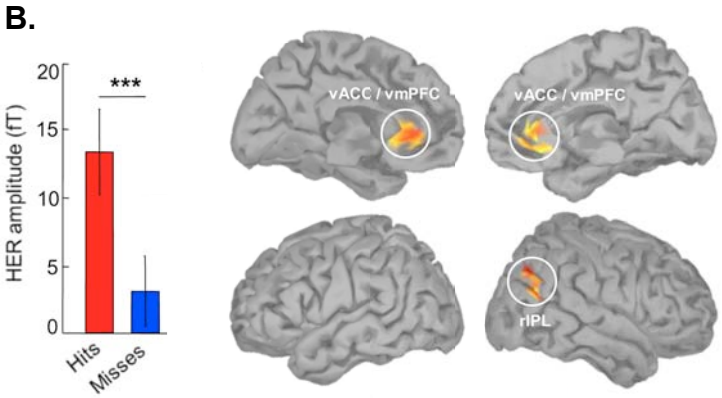
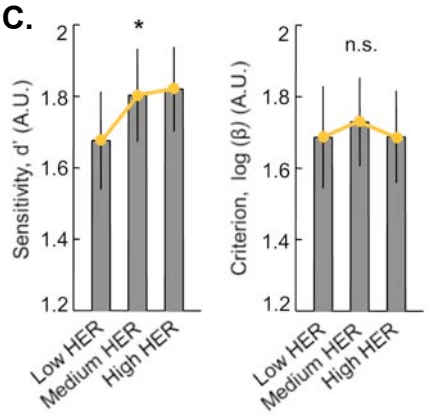
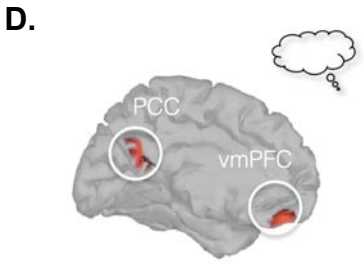

Subjective visual experience
Visual content
Egocentric neural reference frame
Preferential locus of conscious experience
Unconscious & automatic feed-forward processing
Optional attentive focusing on details

Figure 1. Subjective visual experience arises from the integration of visual content with an egocentric reference frame based on the neural monitoring of visceral inputs. Left, schematic representation of information flow in the visual hierarchy. Visual processing begins with an unconscious and automatic wave of feed-forward processing, generating an integrated visual scene representation in higher-order visual areas. Conscious percepts are preferentially formed at this level. The conscious retrieval of details would require an additional and optional descending processing. Modified from (Campana and Tallon-Baudry, 2013). Right, the neural monitoring of ascending visceral inputs creates an egocentric reference frame, from which first-person perspective can develop. Modified from (Park and Tallon-Baudry, 2014). The integration of visual content and the egocentric reference frame gives rise to subjective visual experience.


A.
Fixation
Warning
Stimulus
Delay
Response
HERs in
Hits & Misses
defines
Hits & Misses


B.
HER amplitude (fT)
20
15
10
5
0
***
Hits
Misses
vACC / vmPFC
vACC / vmPFC
rIPL


C.
Sensitivity, d' (A.U.)
2
1.8
1.6
1.4
1.2
*
Low HER
Medium HER
High HER
Criterion, log (β) (A.U.)
n.s.


D.
PCC
vmPFC

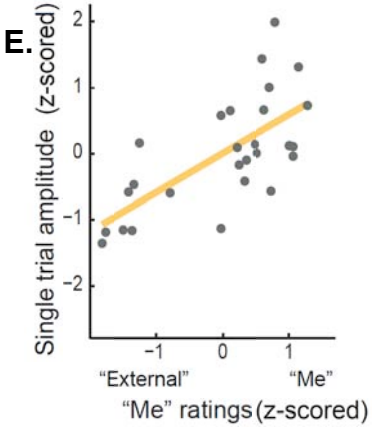

E.
Single trial amplitude (z-scored)
2
1
0
−1
−2
−1
0
1
"External"
"Me"
"Me" ratings (z-scored)

Figure 2. Neural responses to heartbeats and subjectivity. A. Paradigm: participants fixate a central bull's eye that turns red to indicate the beginning of a trial. After a variable delay, a faint stimulus may or may not appear. After another variable delay, participants are prompted to report whether they have seen a stimulus or not. Participants' responses determine hits and misses. Heartbeat-evoked responses (HER) are computed as evoked responses to the heartbeats occurring in the warning interval. **B.** HER amplitude in hits is larger than in misses, in vACC-vmPFC and rIPL. **C.** HER amplitude increase corresponds to a significant increase in perceptual sensitivity (left), while decision criterion does not vary with HER amplitude (right). **D.** Neural responses to heartbeats during spontaneous thoughts co-vary with the self-relatedness of the thought, in the midline nodes of the default network. **E.** Correlation between single thought rating (thought oriented toward an external object or toward oneself) and the amplitude of neural response to heartbeats recorded intracranially from the vmPFC of an epileptic patient. Panels A-C modified from (Park et al., 2014), panel D from (Babo-Rebelo et al., 2016a) and panel E from (Babo-Rebelo et al., 2016b). vACC: ventral anterior cingulate cortex; vmPFC: ventro-medial prefrontal cortex; rIPL: right inferior parietal lobule; PCC: posterior cingulate cortex.

References

Armour JA, Ardell JL (2004) Basic and clinical neurocardiology: Oxford University Press.

Aspell JE, Heydrich L, Marillier G, Lavanchy T, Herbelin B, Blanke O (2013) Turning body and self inside out: visualized heartbeats alter bodily self-consciousness and tactile perception. Psychol Sci 24:2445-2453.

Baars BJ (1997) In the theatre of consciousness: Global workspace theory, a rigorous scientific theory of consciousness. Journal of Consciousness Studies 4:292-309.

Babo-Rebelo M, Richter CG, Tallon-Baudry C (2016a) Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. Journal of Neuroscience in press.

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C (2016b) Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philos T R Soc B 371.

Beissner F, Meissner K, Bar KJ, Napadow V (2013) The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. J Neurosci 33:10503-10511.

Bhatti JA, Nathens AB, Thiruchelvam D, Grantcharov T, Goldstein BI, Redelmeier DA (2016) Self-harm Emergencies After Bariatric Surgery: A Population-Based Cohort Study. JAMA Surg 151:226-232.

Blanke O (2012) Multisensory brain mechanisms of bodily self-consciousness. Nat Rev Neurosci 13:556-571.

Block N (2007) Consciousness, accessibility, and the mesh between psychology and neuroscience. Behav Brain Sci 30:481-499; discussion 499-548.

Buffalo EA, Fries P, Landman R, Liang H, Desimone R (2010) A backward progression of attentional effects in the ventral stream. Proc Natl Acad Sci U S A 107:361-365.

Campana F, Tallon-Baudry C (2013) Anchoring visual subjective experience in a neural model: the coarse vividness hypothesis. Neuropsychologia 51:1050-1060.

Campana F, Rebollo I, Urai A, Wyart V, Tallon-Baudry C (2016) Conscious Vision Proceeds from Global to Local Content in Goal-Directed Tasks and Spontaneous Vision. J Neurosci 36:5200-5213.

Canales-Johnson A, Silva C, Huepe D, Rivera-Rei A, Noreika V, Garcia Mdel C, Silva W, Ciraolo C, Vaucheret E, Sedeno L, Couto B, Kargieman L, Baglivo F, Sigman M, Chennu S, Ibanez A, Rodriguez E, Bekinschtein TA (2015) Auditory Feedback Differentially Modulates Behavioral and Neural Markers of Objective and Subjective Performance When Tapping to Your Heartbeat. Cereb Cortex 25:4490-4503.

Carrasco M, Ling S, Read S (2004) Attention alters appearance. Nat Neurosci 7:308-313.

Chalmers DJ (1995) Facing up to the problem of consciousness. Journal of Consciousness Studies 2:200-219.

Cohen MA, Dennett DC (2011) Consciousness cannot be separated from function. Trends Cogn Sci 15:358-364.

Couto B, Adolfi F, Velasquez M, Mesow M, Feinstein J, Canales-Johnson A, Mikulan E, Martinez-Pernia D, Bekinschtein T, Sigman M, Manes F, Ibanez A (2015) Heart evoked potential triggers brain responses to natural affective scenes: A preliminary study. Auton Neurosci 193:132-137.

Craig AD (2002) How do you feel? Interoception: the sense of the physiological condition of the body. Nat Rev Neurosci 3:655-666.

Craig AD (2009) How do you feel--now? The anterior insula and human awareness. Nat Rev Neurosci 10:59-70.
Critchley HD, Harrison NA (2013) Visceral influences on brain and behavior. Neuron 77:624-638.
Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ (2004) Neural systems supporting interoceptive awareness. Nat Neurosci 7:189-195.
Damasio A (2010) Self comes to mind: Constructing the conscious brain. New York: Pantheon Books.
Damasio AR (1996) The somatic marker hypothesis and the possible functions of the prefrontal cortex. Philos Trans R Soc Lond B Biol Sci 351:1413-1420.
David N, Newen A, Vogeley K (2008) The "sense of agency" and its underlying cognitive and neural mechanisms. Consciousness and Cognition 17:523-534.
Dehaene S, Naccache L (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition 79:1-37.
Dehaene S, Changeux JP, Naccache L, Sackur J, Sergent C (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends Cogn Sci 10:204-211.
Delfini LF, Campos JJ (1972) Signal detection and the "cardiac arousal cycle". Psychophysiology 9:484-491.
Dennett DC (1991) Consciousness explained. Boston, MA: Little, Brown & Co.
Dirlich G, Vogl L, Plaschke M, Strian F (1997) Cardiac field effects on the EEG. Electroencephalogr Clin Neurophysiol 102:307-315.
Dum RP, Levinthal DJ, Strick PL (2009) The spinothalamic system targets motor and sensory areas in the cerebral cortex of monkeys. J Neurosci 29:14223-14235.
Elliott R, Graf V (1972) Visual sensitivity as a function of phase of cardiac cycle. Psychophysiology 9:357-361.
Fan J, Xu P, Van Dam NT, Eilam-Stock T, Gu X, Luo YJ, Hof PR (2012) Spontaneous brain activity relates to autonomic arousal. J Neurosci 32:11176-11186.
Ferre ER, Lopez C, Haggard P (2014) Anchoring the Self to the Body: Vestibular Contribution to the Sense of Self. Psychol Sci.
Frässle S, Sommer J, Jansen A, Naber M, Einhäuser W (2014) Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. Journal of Neuroscience 34:1738-1747.
Fukushima H, Terasawa Y, Umeda S (2011) Association between interoception and empathy: Evidence from heart-beat evoked brain potential. Int J Psychophysiol 79:259-265.
Furness JB (2006) The enteric nervous system: Blackwell Publishing.
Gallagher S (2000) Philosophical conceptions of the self: implications for cognitive science. Trends in Cognitive Sciences 4:14-21.
Henry TR (2002) Therapeutic mechanisms of vagus nerve stimulation. Neurology 59:S3-14.
Hester R, Foxe JJ, Molholm S, Shpaner M, Garavan H (2005) Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. Neuroimage 27:602-608.
Hochstein S, Ahissar M (2002) View from the top: hierarchies and reverse hierarchies in the visual system. Neuron 36:791-804.
Hsu SM, George N, Wyart V, Tallon-Baudry C (2011) Voluntary and involuntary spatial attentions interact differently with awareness. Neuropsychologia 49:2465-2474.
Kelly KA, Code CF (1971) Canine gastric pacemaker. Am J Physiol 220:112-118.

Kentridge RW, Heywood CA, Weiskrantz L (1999) Attention without awareness in blindsight. Proceedings of the Royal Society of London Series B - Biological Sciences 266:1805-1811.
Kentridge RW, Heywood CA, Weiskrantz L (2004) Spatial attention speeds discrimination without awareness in blindsight. Neuropsychologia 42:831-835.
Kern M, Aertsen A, Schulze-Bonhage A, Ball T (2013) Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. Neuroimage 81C:178-190.
Khalsa SS, Rudrauf D, Feinstein JS, Tranel D (2009) The pathways of interoceptive awareness. Nat Neurosci 12:1494-1496.
Kim KJ, Ramiro Diaz J, Iddings JA, Filosa JA (2016) Vasculo-Neuronal Coupling: Retrograde Vascular Communication to Brain Neurons. J Neurosci 36:12624-12639.
Kimchi R (1992) Primacy of wholistic processing and global/local paradigm: a critical review. Psychol Bull 112:24-38.
King JR, Pescetelli N, Dehaene S (2016) Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. Neuron 92:1122-1134.
Koivisto M, Revonsuo A, Lehtonen M (2006) Independence of visual awareness from the scope of attention: an electrophysiological study. Cereb Cortex 16:415-424.
Lafer-Sousa R, Hermann KL, Conway BR (2015) Striking individual differences in color perception uncovered by 'the dress' photograph. Curr Biol 25:R545-546.
Lau HC, Passingham RE (2007) Unconscious activation of the cognitive control system in the human prefrontal cortex. J Neurosci 27:5805-5811.
Lechinger J, Heib DP, Gruber W, Schabus M, Klimesch W (2015) Heartbeat-related EEG amplitude and phase modulations from wakefulness to deep sleep: Interactions with sleep spindles and slow oscillations. Psychophysiology 52:1441-1450.
Luck SJ, Vogel EK, Shapiro KL (1996) Word meaning can be accessed but not reported during the attentional blink. Nature 383:616-618.
Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol 77:24-42.
Luft CD, Bhattacharya J (2015) Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. Sci Rep 5:15717.
Mack A, Rock I (1998) Inattentional Blindness. Cambridge, Massachussets: MIT Press.
Mehta SD, Ulbert I, Schroeder CE (2000) Intermodal selective attention in monkeys. I: Distribution and timing of effects across visual areas. Cereb Cortex 10:343-358.
Montoya P, Schandry R, Muller A (1993) Heartbeat evoked potentials (HEP): topography and influence of cardiac awareness and focus of attention. Electroencephalogr Clin Neurophysiol 88:163-172.
Murphy PR, O'Connell RG, O'Sullivan M, Robertson IH, Balsters JH (2014) Pupil diameter covaries with BOLD activity in human locus coeruleus. Hum Brain Mapp 35:4140-4154.
Nagai Y, Critchley HD, Featherstone E, Trimble MR, Dolan RJ (2004) Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: a physiological account of a "default mode" of brain function. Neuroimage 22:243-251.
Nagel T (1974) What is it like to be a bat? The Philosophical Review 83:435-450.
Navon D (1977) Forest before trees: the precedence of global features in visual perception. Cognitive Psychology 9:353-383.

Norman LJ, Heywood CA, Kentridge RW (2013) Object-Based Attention Without Awareness. Psychological Science 24:836-843.
O'Regan JK, Noe A (2001) A sensorimotor account of vision and visual consciousness. Behav Brain Sci 24:939-973; discussion 973-1031.
Ohl S, Wohltat C, Kliegl R, Pollatos O, Engbert R (2016) Microsaccades Are Coupled to Heartbeat. J Neurosci 36:1237-1241.
Park HD, Tallon-Baudry C (2014) The neural subjective frame: from bodily signals to perceptual consciousness. Philos Trans R Soc Lond B Biol Sci 369:20130208.
Park HD, Correia S, Ducorps A, Tallon-Baudry C (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat Neurosci 17:612-618.
Park HD, Bernasconi F, Bello-Ruiz J, Pfeiffer C, Salomon R, Blanke O (2016) Transient Modulations of Neural Responses to Heartbeats Covary with Bodily Self-Consciousness. J Neurosci 36:8453-8460.
Petkova VI, Bjornsdotter M, Gentile G, Jonsson T, Li TQ, Ehrsson HH (2011) From part- to whole-body ownership in the multisensory brain. Curr Biol 21:1118-1122.
Pollatos O, Schandry R (2004) Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. Psychophysiology 41:476-482.
Pritchard TC, Hamilton RB, Norgren R (2000) Projections of the parabrachial nucleus in the old world monkey. Exp Neurol 165:101-117.
Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? Neuroimage 57:1221-1233.
Rensink RA (2002) Change detection. Annu Rev Psychol 53:245-277.
Richter CG, Babo-Rebelo M, Schwartz D, Tallon-Baudry C (2017) Phase-amplitude coupling at the organism level: The amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm. NeuroImage 146:951-958.
Salomon R, Ronchi R, Donz J, Bello-Ruiz J, Herbelin B, Martet R, Faivre N, Schaller K, Blanke O (2016) The Insula Mediates Access to Awareness of Visual Stimuli Presented Synchronously to the Heartbeat. J Neurosci 36:5115-5127.
Sanders KM, Koh SD, Ward SM (2006) Interstitial cells of Cajal as pacemakers in the gastrointestinal tract. Annu Rev Physiol 68:307-343.
Schandry R, Montoya P (1996) Event-related brain potentials and the processing of cardiac activity. Biol Psychol 42:75-85.
Schandry R, Sparrer B, Weitkunat R (1986) From the heart to the brain: a study of heartbeat contingent scalp potentials. Int J Neurosci 30:261-275.
Schurger A, Cowey A, Cohen JD, Treisman A, Tallon-Baudry C (2008) Distinct and independent correlates of attention and awareness in a hemianopic patient. Neuropsychologia 46:2189-2197.
Searle JR (2000) Consciousness. Annu Rev Neurosci 23:557-578.
Sel A, Azevedo RT, Tsakiris M (in press) Heartfelt Self: Cardio-Visual Integration Affects Self-Face Recognition and Interoceptive Cortical Processing. Cerebral Cortex.
Sergent C, Wyart V, Babo-Rebelo M, Cohen L, Naccache L, Tallon-Baudry C (2013) Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. Curr Biol 23:150-155.
Shapiro KL, Raymond JE, Arnell KM (1997) The attentional blink. Trends Cogn Sci 1:291-296.
Shepherd JT (1985) The heart as a sensory organ. J Am Coll Cardiol 5:83B-87B.
Simons DJ, Levin DT (1997) Change blindness. Trends Cogn Sci 1:261-267.

Solomon JA (2004) The effect of spatial cues on visual sensitivity. Vision Res 44:1209-1216.
Soto D, Mantyla T, Silvanto J (2011) Working memory without consciousness. Curr Biol 21:R912-913.
Sumner P, Nachev P, Morris P, Peters AM, Jackson SR, Kennard C, Husain M (2007) Human medial frontal cortex mediates unconscious inhibition of voluntary action. Neuron 54:697-711.
Suzuki K, Garfinkel SN, Critchley HD, Seth AK (2013) Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. Neuropsychologia 51:2909-2917.
Tallon-Baudry C (2012) On the neural mechanisms subserving consciousness and attention. Front Psychology 2:397.
Thayer JF, Ahs F, Fredrikson M, Sollers JJ, 3rd, Wager TD (2012) A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neurosci Biobehav Rev 36:747-756.
Tindle HA, Omalu B, Courcoulas A, Marcus M, Hammers J, Kuller LH (2010) Risk of suicide after long-term follow-up from bariatric surgery. Am J Med 123:1036-1042.
Tononi G, Koch C (2008) The neural correlates of consciousness - An update. Ann N Y Acad Sci 1124:239-261.
van Boxtel JJ, Tsuchiya N, Koch C (2010) Opposing effects of attention and consciousness on afterimages. Proc Natl Acad Sci U S A 107:8883-8888.
Van der Werf YD, Witter MP, Groenewegen HJ (2002) The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. Brain Res Brain Res Rev 39:107-140.
van Gaal S, Ridderinkhof KR, Fahrenfort JJ, Scholte HS, Lamme VA (2008) Frontal cortex mediates unconsciously triggered inhibitory control. J Neurosci 28:8053-8062.
Velden M, Juris M (1975) Perceptual performance as a function of intra-cycle cardiac activity. Psychophysiology 12:685-692.
Vogt BA, Pandya DN, Rosene DL (1987) Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. The Journal of comparative neurology 262:256-270.
Watanabe M, Cheng K, Murayama Y, Ueno K, Asamizuya T, Tanaka K, Logothetis N (2011) Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. Science 334:829-831.
Webb TW, Igelström KM, Schurger A, Graziano M (2017) Cortical networks involved in visual awareness independent of visual attention. Proc Natl Acad Sci U S A 113:13923-13928.
Wong SW, Masse N, Kimmerly DS, Menon RS, Shoemaker JK (2007) Ventral medial prefrontal cortex and cardiovagal control in conscious humans. Neuroimage 35:698-708.
Wyart V, Tallon-Baudry C (2008) A neural dissociation between visual awareness and spatial attention. J Neurosci 28:2667-2679.
Wyart V, Dehaene S, Tallon-Baudry C (2012) Early dissociation between neural signatures of endogenous spatial attention and perceptual awareness during visual masking. Front Hum Neurosci 6.
Zahavi D (2003) Husserl's phenomenology. Stanford, California: Stanford University Press.

References

Adler D, Herbelin B, Similowski T, Blanke O. 2014. Breathing and sense of self: Visuo–respiratory conflicts alter body self-consciousness. *Respir. Physiol. Neurobiol.* 203:68–74

Adolfi F, Couto B, Richter F, Decety J, Lopez J, et al. 2017. Convergence of interoception, emotion, and social cognition: A twofold fMRI meta-analysis and lesion approach. *Cortex*. 88:124–42

Ainley V, Apps MAJ, Fotopoulou A, Tsakiris M. 2016. "Bodily precision": a predictive coding account of individual differences in interoceptive accuracy. *Philos. Trans. R. Soc. B.* 371(1708):1–9

Amsterdam B. 1972. Mirror self-image reactions before age two. *Dev. Psychobiol.* 5(4):297–305

Andrews-Hanna JR, Kaiser RH, Turner AEJ, Reineberg AE, Godinez D, et al. 2013. A penny for your thoughts: dimensions of self-generated thought content and relationships with individual differences in emotional wellbeing. *Front. Psychol.* 4(November):1–13

Andrews-Hanna JR, Reidler JS, Huang C, Buckner RL. 2010a. Evidence for the default network's role in spontaneous cognition. *J. Neurophysiol.* 104(1):322–35

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL. 2010b. Functional-anatomic fractionation of the brain's default network. *Neuron*. 65(4):550–62

Andrews-Hanna JR, Smallwood J, Spreng RN. 2014. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N. Y. Acad. Sci.* 1316:29–52

Apps MAJ, Tajadura-Jiménez A, Turley G, Tsakiris M. 2012. The different faces of one 's self : an fMRI study into the recognition of current and past self-facial appearances. *Neuroimage*. 63(3):1720–29

Araujo HF, Kaplan J, Damasio A. 2013. Cortical midline structures and autobiographical-self processes: an activation-likelihood estimation meta-analysis. *Front. Hum. Neurosci.* 7:1–10

Armour JA, Ardell JL. 2004. *Basic and Clinical Neurocardiology*. New York: Oxford University Press

Aspell JE, Heydrich L, Marillier G, Lavanchy T, Herbelin B, Blanke O. 2013. Turning body and self inside out: visualized heartbeats alter bodily self-consciousness and tactile

perception. *Psychol. Sci.* 24(12):2445–53

Avery JA, Kerr KL, Ingeholm JE, Burrows K, Bodurka J, Simmons WK. 2015. A common gustatory and interoceptive representation in the human mid-insula. *Hum. Brain Mapp.* 36(8):2996–3006

Azevedo RT, Garfinkel SN, Critchley HD, Tsakiris M. 2017. Cardiac afferent activity modulates the expression of racial stereotypes. *Nat. Commun.* 8(13854):1–9

Babo-Rebelo M, Richter C, Tallon-Baudry C. 2016a. Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. *J. Neurosci.* 36(30):7829–40

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C. 2016b. Is the cardiac monitoring function related to the self in both the default network and right anterior insula? *Philos. Trans. R. Soc. B.* 371(1708):1–13

Beckmann M, Johansen-Berg H, Rushworth MFS. 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J. Neurosci.* 29(4):1175–90

Beissner F, Meissner K, Bär K-J, Napadow V. 2013. The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. *J. Neurosci.* 33(25):10503–11

Bergouignan L, Nyberg L, Ehrsson HH. 2014. Out-of-body-induced hippocampal amnesia. *Proc Natl Acad Sci USA.* 111(12):4421–26

Bernier PM, Grafton ST. 2010. Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. *Neuron.* 68(4):776–88

Berscheid E, Snyder M, Omoto AM. 1989. The relationship closeness inventory : Assessing the closeness of interpersonal relationships. *J. Pers. Soc. Psychol.* 57(5):792–807

Birn RM. 2012. The role of physiological noise in resting-state functional connectivity. *Neuroimage.* 62(2):864–70

Blanke O. 2005. Linking Out-of-Body Experience and Self Processing to Mental Own-Body Imagery at the Temporoparietal Junction. *J. Neurosci.* 25(3):550–57

Blanke O. 2012. Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* 13(8):556–71

Blanke O, Metzinger T. 2009. Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.* 13(1):7–13

Blanke O, Ortigue S, Landis T, Seeck M. 2002. Stimulating illusory own-body perceptions. *Nature.* 419(6904):269–70

Blanke O, Slater M, Serino A. 2015. Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron*. 88(1):145–66

Boccia M, Nemmi F, Guariglia C. 2014. Neuropsychology of environmental navigation in humans: Review and meta-analysis of fMRI studies in healthy participants. *Neuropsychol. Rev.* 24(4):236–51

Botvinick M, Cohen J. 1998. Rubber hands "feel" touch that eyes see. *Nature*. 391(6669):756

Brener J, Ring C. 2016. Towards a psychophysics of interoceptive processes: the measurement of heartbeat detection. *Philos. Trans. R. Soc. B*. 371(1708):1–9

Bruno M-A, Bernheim JL, Ledoux D, Pellas F, Demertzi A, Laureys S. 2011. A survey on self-assessed well-being in a cohort of chronic locked-in syndrome patients: happy majority, miserable minority. *BMJ Open*. 1(1):e000039

Buckner RL, Andrews-Hanna JR, Schacter DL. 2008. The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124:1–38

Burgess N, Becker S, King JA, O'Keefe J. 2001. Memory for events and their spatial context: models and experiments. *Philos. Trans. R. Soc. L. B Biol Sci.* 356(1413):1493–1503

Bzdok D, Heeger A, Langner R, Laird AR, Fox PT, et al. 2015. Subspecialization in the human posterior medial cortex. *Neuroimage*. 106:55–71

Bzdok D, Langner R, Schilbach L, Jakobs O, Roski C, et al. 2013. Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. *Neuroimage*. 81:381–92

Cabeza R, St Jacques P. 2007. Functional neuroimaging of autobiographical memory. *Trends Cogn. Sci.* 11(5):219–27

Canales-Johnson A, Silva C, Huepe D, Rivera-Rei A, Noreika V, et al. 2015. Auditory feedback differentially modulates behavioral and neural markers of objective and subjective performance when tapping to your heartbeat. *Cereb. Cortex*. 25(11):4490–4503

Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K. 2006. The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. *Neuroimage*. 33(2):430–48

Cavanna AE, Trimble MR. 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*. 129(Pt 3):564–83

Cavazzana A, Penolazzi B, Begliomini C, Bisiacchi PS. 2015. Neural underpinnings of the "agent brain": New evidence from transcranial direct current stimulation. *Eur. J. Neurosci.* 42(3):1889–94

Cechetto DF, Saper CB. 1987. Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. *J. Comp. Neurol.* 262(1):27–45

Chambon V, Wenke D, Fleming SM, Prinz W, Haggard P. 2013. An online neural substrate for sense of agency. *Cereb. Cortex.* 23(5):1031–37

Chang C, Metzger CD, Glover GH, Duyn JH, Heinze H-J, Walter M. 2013. Association between heart rate variability and fluctuations in resting-state functional connectivity. *Neuroimage.* 68:93–104

Christoff K, Cosmelli D, Legrand D, Thompson E. 2011. Specifying the self for cognitive neuroscience. *Trends Cogn. Sci.* 15(3):104–12

Christoff K, Gordon AM, Smallwood J, Smith R, Schooler JW. 2009. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc Natl Acad Sci USA.* 106(21):8719–24

Christoff K, Irving ZC, Fox KCR, Spreng RN, Andrews-Hanna JR. 2016. Mind-wandering as spontaneous thought: a dynamic framework. *Nat. Rev. Neurosci.* 17(11):718–31

Ciaramelli E, Rosenbaum RS, Solcz S, Levine B, Moscovitch M. 2010. Mental space travel: damage to posterior parietal cortex prevents egocentric navigation and reexperiencing of remote spatial memories. *J. Exp. Psychol. Learn. Mem. Cogn.* 36(3):619–34

Coleman MN, Ross CF. 2004. Primate auditory diversity and its influence on hearing performance. *Anat. Rec. - Part A Discov. Mol. Cell. Evol. Biol.* 281(1):1123–37

Couto B, Adolfi F, Velasquez M, Mesow M, Feinstein J, et al. 2015. Heart evoked potential triggers brain responses to natural affective scenes: A preliminary study. *Auton. Neurosci. Basic Clin.* 193:132–37

Couto B, Salles A, Sedeño L, Peradejordi M, Barttfeld P, et al. 2014. The man who feels two hearts: the different pathways of interoception. *Soc. Cogn. Affect. Neurosci.* 9(9):1253–60

Craig AD. 2002. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3(8):655–66

Craig AD. 2009. How do you feel - now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10(1):59–70

Critchley HD, Harrison NA. 2013. Visceral influences on brain and behavior. *Neuron.* 77(4):624–38

Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ. 2004. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7(2):189–95

D'Argembeau A. 2013. On the role of the ventromedial prefrontal cortex in self-processing: the valuation hypothesis. *Front. Hum. Neurosci.* 7:372

D'Argembeau A, Collette F, Van der Linden M, Laureys S, Del Fiore G, et al. 2005. Self-referential reflective activity and its relationship with rest: a PET study. *Neuroimage*. 25(2):616–24

D'Argembeau A, Feyers D, Majerus S, Collette F, Van der Linden M, et al. 2008. Self-reflection across time: cortical midline structures differentiate between present and past selves. *Soc. Cogn. Affect. Neurosci.* 3(3):244–52

Damasio A. 1994. *Descartes' Error. Emotion, Reason, and the Human Brain*. New York: Avon Books

Damasio A, Carvalho GB. 2013. The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* 14(2):143–52

Damasio AR. 1999. *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. New York: Harcourt

Damasio AR. 2003. *Looking for Spinoza. Joy, Sorrow, and the Feeling Brain*. Harvest

Damasio AR, Damasio H, Tranel D. 2012. Persistence of feelings and sentience after bilateral damage of the insula. *Cereb. cortex.* 23(4):833–46

Daprati E, Franck N, Georgieff N, Proust J, Pacherie E, et al. 1997. Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*. 65(1):71–86

Davey CG, Pujol J, Harrison BJ. 2016. Mapping the self in the brain's default mode network. *Neuroimage*. 132:390–97

David N, Stenzel A, Schneider TR, Engel AK. 2011. The feeling of agency: empirical indicators for a pre-reflective level of action awareness. *Front. Psychol.* 2:149

de Munck JC, Gonçalves SI, Faes TJC, Kuijer JPA, Pouwels PJW, et al. 2008. A study of the brain's resting state based on alpha band power, heart rate and fMRI. *Neuroimage*. 42(1):112–21

de Pasquale F, Della Penna S, Snyder AZ, Lewis C, Mantini D, et al. 2010. Temporal dynamics of spontaneous MEG activity in brain networks. *Proc Natl Acad Sci USA*. 107(13):6040–45

De Ridder D, Van Laere K, Dupont P, Menovsky T, Van de Heyning P. 2007. Visualizing out-of-body experience in the brain. *N. Engl. J. Med.* 357(18):1829–33

de Vignemont F. 2011. Embodiment, ownership and disownership. *Conscious. Cogn.*

20(1):82–93

Deen B, Pitskel NB, Pelphrey K a. 2011. Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex.* 21(7):1498–1506

Delamillieure P, Doucet G, Mazoyer B, Turbelin M-R, Delcroix N, et al. 2010. The resting state questionnaire: An introspective questionnaire for evaluation of inner experience during the conscious resting state. *Brain Res. Bull.* 81(6):565–73

Dennett D. 1991. *Consciousness Explained*. Little Brown & Co

Denny BT, Kober H, Wager TD, Ochsner KN. 2012. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* 24(8):1742–52

Devue C, Collette F, Balteau E, Degueldre C, Luxen A, et al. 2007. Here I am: the cortical correlates of visual self-recognition. *Brain Res.* 1143:169–82

Dijkstra K, Kaschak MP, Zwaan RA. 2007. Body posture facilitates retrieval of autobiographical memories. *Cognition.* 102(1):139–49

Dijkstra N, Bosch S, van Gerven MAJ. 2017. Vividness of visual imagery depends on the neural overlap with perception in visual areas. *J. Neurosci.* 37(5):1367–73

Dirlich G, Dietl T, Vogl L, Strian F. 1998. Topography and morphology of heart action-related EEG potentials. *Electroencephalogr. Clin. Neurophysiol.* 108(3):299–305

Dirlich G, Vogl L, Plaschke M, Strian F. 1997. Cardiac field effects on the EEG. *Electroencephalogr. Clin. Neurophysiol.* 102(4):307–15

Driver ID, Whittaker JR, Bright MG, Muthukumaraswamy SD, Murphy K. 2016. Arterial CO 2 fluctuations modulate neuronal rhythmicity : implications for MEG and fMRI studies of resting-state networks. *J. Neurosci.* 36(33):8541–50

Dubois D, Ameis SH, Lai M, Casanova M, Desarkar P. 2016. Interoception in Autism Spectrum Disorder : A Systematic Review. *Int. J. Dev. Neurosci.* 52:104–11

Dum RP, Levinthal DJ, Strick PL. 2009. The spinothalamic system targets motor and sensory areas in the cerebral cortex of monkeys. *J. Neurosci.* 29(45):14223–35

Dunn BD, Galton HC, Morgan R, Evans D, Oliver C, et al. 2010. Listening to your heart. How interoception shapes emotion experience and intuitive decision making. *Psychol. Sci.* 21(12):1835–44

Edwards L, McIntyre D, Carroll D, Ring C, Martin U. 2002. The human nociceptive flexion reflex threshold is higher during systole than diastole. *Psychophysiology.* 39(5):678–81

Edwards L, Ring C, McIntyre D, Carroll D, Martin U. 2007. Psychomotor speed in hypertension: Effects of reaction time components, stimulus modality, and phase of the cardiac cycle. *Psychophysiology*. 44(3):459–68

Ehrsson HH. 2007. The experimental induction of out-of-body experiences. *Science*. 317(5841):1048

Ehrsson HH, Holmes NP, Passingham RE. 2005. Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. *J. Neurosci.* 25(45):10564–73

Ehrsson HH, Spence C, Passingham RE. 2004. That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Sci. (New York, NY)*. 305(5685):875–77

Ehrsson HH, Wiech K, Weiskopf N, Dolan RJ, Passingham RE. 2007. Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proc. Natl. Acad. Sci. U. S. A.* 104(23):9828–33

Epstein R, Lanza RP, Skinner BF. 1981. "Self-Awareness" in the Pigeon. *Science*. 212(4495):695–96

Fagius J, Wallin BG. 1980. Sympathetic reflex latencies and conduction velocities in normal man. *J. Neurol. Sci.* 47(3):433–48

Fan J, Xu P, Van Dam NT, Eilam-Stock T, Gu X, et al. 2012. Spontaneous brain activity relates to autonomic arousal. *J. Neurosci.* 32(33):11176–86

Farb NAS, Segal Z V, Anderson AK. 2012. Attentional Modulation of Primary Interoceptive and Exteroceptive Cortices. *Cereb. cortex*. 1–13

Farrer C, Franck N, Georgieff N, Frith C., Decety J, Jeannerod M. 2003. Modulating the experience of agency: a positron emission tomography study. *Neuroimage*. 18(2):324–33

Farrer C, Frith CD. 2002. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage*. 15(3):596–603

Fischl B, Van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, et al. 2004. Automatically parcellating the human cerebral cortex. *Cereb. Cortex*. 14(1):11–22

Fossati P. 2013. Imaging autobiographical memory. *Dialogues Clin. Neurosci.* 15(4):487–90

Fotopoulou A, Jenkinson PM, Tsakiris M, Haggard P, Rudd A, Kopelman MD. 2011. Mirror-view reverses somatoparaphrenia: Dissociation between first- and third-person perspectives on body ownership. *Neuropsychologia*. 49(14):3946–55

Fox KCRR, Spreng RN, Ellamil M, Andrews-Hanna JR, Christoff K. 2015. The wandering

brain: meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage*. 111:611–21

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA*. 102(27):9673–78

Fransson P, Marrelec G. 2008. The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *Neuroimage*. 42(3):1178–84

Frässle S, Sommer J, Jansen A, Naber M, Einhauser W. 2014. Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J. Neurosci.* 34(5):1738–47

Freton M, Lemogne C, Bergouignan L, Delaveau P, Lehéricy S, Fossati P. 2014. The eye of the self: precuneus volume and visual perspective during autobiographical memory retrieval. *Brain Struct. Funct.* 219(3):959–68

Frith CD, Blakemore SJ, Wolpert DM. 2000. Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 355(1404):1771–88

Fukushima H, Terasawa Y, Umeda S. 2011. Association between interoception and empathy: evidence from heartbeat-evoked brain potential. *Int. J. Psychophysiol.* 79(2):259–65

Gallagher S. 2000. Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4(1):14–21

Gallagher S. 2012. *Phenomenology*. Palgrave Macmillan

Gallagher S. 2013. A pattern theory of self. *Front. Hum. Neurosci.* 7(August):443

Gallagher S, Zahavi D. 2008. *The Phenomenological Mind*

Gallup GG. 1970. Chimpanzees: self recognition. *Science*. 167(3914):86–87

Gallup GG, Platek SM, Spaulding KN. 2014. The nature of visual self-recognition revisited. *Trends Cogn. Sci.* 18(2):57–58

García-Cordero I, Sedeño L, De La Fuente L, Slachevsky A, Forno G, et al. 2016. Feeling, learning from, and being aware of inner states: Interoceptive dimensions in neurodegeneration and stroke. *Philos. Trans. R. Soc. London B Biol. Sci.* 371(1708):1–10

Garfinkel SN, Barrett AB, Minati L, Dolan RJ, Seth AK, Critchley HD. 2013a. What the heart forgets: Cardiac timing influences memory for words and is modulated by metacognition and interoceptive sensitivity. *Psychophysiology*. 50(6):505–12

Garfinkel SN, Critchley HD. 2015. Threat and the body: How the heart supports fear processing. *Trends Cogn. Sci.* 20(1):34–46

Garfinkel SN, Minati L, Gray MA, Seth AK, Dolan RJ, Critchley HD. 2014. Fear from the heart: Sensitivity to fear stimuli depends on individual heartbeats. *J. Neurosci.* 34(19):6573–82

Garfinkel SN, Nagai Y, Seth AK, Critchley HD. 2013b. Neuroimaging studies of interoception and self-awareness. In *Neuroimaging of Consciousness*, eds. AE Cavanna, A Nani, H Blumenfeld, S Laureys, pp. 207–24. Springer Berlin Heidelberg

Garfinkel SN, Seth AK, Barrett AB, Suzuki K, Critchley HD. 2015. Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biol. Psychol.* 104:65–74

Giambra LM. 1993. The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Exp. Gerontol.* 28(4–5):485–92

Gillihan SJ, Farah MJ. 2005. Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychol. Bull.* 131(1):76–97

Glover GH, Li TQ, Ress D. 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44(1):162–67

Goldberg II, Harel M, Malach R. 2006. When the brain loses its self: prefrontal inactivation during sensorimotor processing. *Neuron.* 50(2):329–39

Golland Y, Bentin S, Gelbard H, Benjamini Y, Heller R, et al. 2007. Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cereb. cortex.* 17(4):766–77

Golland Y, Golland P, Bentin S, Malach R. 2008. Data-driven clustering reveals a fundamental subdivision of the human cortex into two global systems. *Neuropsychologia.* 46(2):540–53

Gramann K, Onton J, Riccobon D, Mueller HJ, Bardins S, Makeig S. 2010. Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *J. Cogn. Neurosci.* 22(12):2836–49

Gray MA, Rylander K, Harrison NA, Wallin BG, Critchley HD. 2009. Following one's heart: cardiac rhythms gate central initiation of sympathetic reflexes. *J. Neurosci.* 29(6):1817–25

Gray MA, Taggart P, Sutton PM, Groves D, Holdright DR, et al. 2007. A cortical potential reflecting cardiac function. *Proc Natl Acad Sci USA.* 104(16):6818–23

Gusnard DA, Akbudak E, Shulman GL, Raichle ME. 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci USA*. 98(7):4259–64

Guterstam A, Bjornsdotter M, Gentile G, Ehrsson HH. 2015. Posterior Cingulate Cortex Integrates the Senses of Self-Location and Body Ownership. *Curr. Biol.* 25(11):1416–25

Haggard P. 2017. Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18:196–207

Haggard P, Clark S, Kalogeras J. 2002. Voluntary action and conscious awareness. *Nat. Neurosci.* 5(4):382–85

Harver A, Katkin ES, Bloch E. 1993. Signal-detection outcomes on heartbeat and respiratory resistance detection tasks in male and female subjects. *Psychophysiology*. 30(3):223–30

Herbert BM, Muth ER, Pollatos O, Herbert C. 2012. Interoception across modalities: on the relationship between cardiac awareness and the sensitivity for gastric functions. *PLoS One*. 7(5):e36646

Herbert BM, Pollatos O. 2012. The body in the mind: on the relationship between interoception and embodiment. *Top. Cogn. Sci.* 4(4):692–704

Herbert BM, Pollatos O, Schandry R. 2007. Interoceptive sensitivity and emotion processing: an EEG study. *Int. J. Psychophysiol.* 65(3):214–27

Hu C, Di X, Eickhoff SB, Zhang M, Peng K, et al. 2016. Distinct and common aspects of physical and psychological self-representation in the brain: A meta-analysis of self-bias in facial and self-referential judgements. *Neurosci. Biobehav. Rev.* 61:197–207

Huang R-S, Sereno MI. 2013. Bottom-up retinotopic organization supports top-down mental imagery. *Open Neuroimag. J.* 7(858):58–67

Hurlburt RT, Heavey CL. 2001. Telling what we know: describing inner experience. *Trends Cogn. Sci.* 5(9):400–403

Iacovella V, Hasson U. 2011. The relationship between BOLD signal and autonomic nervous system functions: implications for processing of "physiological noise." *Magn. Reson. Imaging*. 29(10):1338–45

Imafuku M, Hakuno Y, Uchida-Ota M, Yamamoto J-I, Minagawa Y. 2014. "Mom called me!" Behavioral and prefrontal responses of infants to self-names spoken by their mothers. *Neuroimage*. 103:476–84

Immanuel SA, Pamula Y, Kohler M, Martin J, Kennedy D, et al. 2014. Heartbeat evoked potentials during sleep and daytime behavior in children with sleep-disordered

breathing. *Am. J. Respir. Crit. Care Med.* 190(10):1149–57

Ingvar DH, Schwartz MS. 1974. Blood flow patterns induced in the dominant hemisphere by speech and reading. *Brain*. 97:273–88

Ionta S, Gassert R, Blanke O. 2011a. Multi-sensory and sensorimotor foundation of bodily self-consciousness - an interdisciplinary approach. *Front. Psychol.* 2:383

Ionta S, Heydrich L, Lenggenhager B, Mouthon M, Fornari E, et al. 2011b. Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron*. 70(2):363–74

Ionta S, Martuzzi R, Salomon R, Blanke O. 2014. The brain network reflecting bodily self-consciousness: a functional connectivity study. *Soc. Cogn. Affect. Neurosci.* 70(2):363–74

James C, Henderson L, Macefield VG. 2013. Real-time imaging of brain areas involved in the generation of spontaneous skin sympathetic nerve activity at rest. *Neuroimage*. 74:188–94

James W. 1884. What is an emotion? *Mind*. 9(34):188–205

James W. 1890. *The Principles of Psychology*. NY, US: Henry Holt and Company

Jänig W. 1996. Neurobiology of visceral afferent neurons: Neuroanatomy, functions, organ regulations and sensations. *Biol. Psychol.* 42(1–2):29–51

Jenkins AC, Macrae CN, Mitchell JP. 2008. Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proc Natl Acad Sci USA*. 105(11):4507–12

Jenkins AC, Mitchell JP. 2011. Medial prefrontal cortex subserves diverse forms of self-reflection. *Soc. Neurosci.* 6(3):211–18

Jousmäki V, Hari R. 1996. Cardiac artifacts in magnetoencephalogram. *J. Clin. Neurophysiol.* 13(2):172–76

Jung Kim K, Ramiro Diaz J, Iddings JA, Filosa JA. 2016. Vasculo-neuronal coupling: retrograde vascular communication to brain neurons. *J. Neurosci.* 36(50):12624–39

Kalckert A, Ehrsson HH. 2014. The moving rubber hand illusion revisited: Comparing movements and visuotactile stimulation to induce illusory ownership. *Conscious. Cogn.* 26(1):117–32

Kammers MPM, de Vignemont F, Verhagen L, Dijkerman HC. 2009. The rubber hand illusion in action. *Neuropsychologia*. 47(1):204–11

Kampe KKW, Frith CD, Frith U. 2003. “Hey John”: Signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *J. Neurosci.* 23(12):5258–63

Kandasamy N, Garfinkel SN, Page L, Hardy B, Critchley HD, et al. 2016. Interoceptive ability predicts survival on a London trading floor. *Sci. Rep.* 6:32986

Kaplan JT, Aziz-Zadeh L, Uddin LQ, Iacoboni M. 2008. The self across the senses: an fMRI study of self-face and self-voice recognition. *Soc. Cogn. Affect. Neurosci.* 3(3):218–23

Kass RE, Raftery AE. 1995. Bayes Factors. *J. Am. Stat. Assoc.* 90(430):773–95

Keenan JP, Nelson A, O’Connor M, Pascual-Leone A. 2001. Self-recognition and the right hemisphere. *Nature*. 409(6818):305

Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, Heatherton TF. 2002. Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14(5):785–94

Kern M, Aertsen A, Schulze-Bonhage A, Ball T. 2013. Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. *Neuroimage*. 81:178–90

Khalsa SS, Rudrauf D, Damasio AR, Davidson RJ, Lutz A, Tranel D. 2008. Interoceptive awareness in experienced meditators. *Psychophysiology*. 45(4):671–77

Khalsa SS, Rudrauf D, Feinstein JS, Tranel D. 2009. The pathways of interoceptive awareness. *Nat. Neurosci.* 12(12):1494–96

Killingsworth MA, Gilbert DT. 2010. A wandering mind is an unhappy mind. *Science*. 330:932

Kim H. 2012. A dual-subsystem model of the brain’s default network: Self-referential processing, memory retrieval processes, and autobiographical memory retrieval. *Neuroimage*. 61(4):966–77

Kjaer TW, Nowak M, Lou HC. 2002. Reflective self-awareness and conscious states: PET evidence for a common midline parietofrontal core. *Neuroimage*. 17(2):1080–86

Kleckner IR, Wormwood JB, Simmons WK, Barrett LF, Quigley KS. 2015. Methodological recommendations for a heartbeat detection-based measure of interoceptive sensitivity. *Psychophysiology*. 52(11):1432–40

Klein SB. 2004. The cognitive neuroscience of knowing one’s self. In *The Cognitive Neurosciences*, ed. MS Gazzaniga, pp. 1007–89. Cambridge: MIT Press. 3rd ed.

Koch A, Pollatos O. 2014. Cardiac sensitivity in children: Sex differences and its relationship to parameters of emotional processing. *Psychophysiology*. 51(9):932–41

Kuehn E, Mueller K, Lohmann G, Schuetz-Bosbach S. 2016. Interoceptive awareness changes the posterior insula functional connectivity profile. *Brain Struct. Funct.* 221(3):1555–71

Kühn S, Brass M, Haggard P. 2013. Feeling in control: Neural correlates of experience of agency. *Cortex.* 49(7):1935–42

Laor N, Wolmer L, Wiener Z, Sharon O, Weizman R, et al. 1999. Image vividness as a psychophysiological regulator in posttraumatic stress disorder. *J. Clin. Exp. Neuropsychol.* 21(1):39–48

Lechinger J, Heib DPJ, Gruber W, Schabus M, Klimesch W. 2015. Heartbeat-related EEG amplitude and phase modulations from wakefulness to deep sleep: Interactions with sleep spindles and slow oscillations. *Psychophysiology.* 52(11):1441–50

Leech R, Braga R, Sharp DJ. 2012. Echoes of the brain within the posterior cingulate cortex. *J. Neurosci.* 32(1):215–22

Legrand D. 2007. Pre-reflective self-as-subject from experiential and empirical perspectives. *Conscious. Cogn.* 16(3):583–99

Legrand D, Ruby P. 2009. What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychol. Rev.* 116(1):252–82

Lenggenhager B, Mouthon M, Blanke O. 2009. Spatial aspects of bodily self-consciousness. *Conscious. Cogn.* 18(1):110–17

Lenggenhager B, Tadi T, Metzinger T, Blanke O. 2007. Video ergo sum: manipulating bodily self-consciousness. *Science.* 317(5841):1096–99

Leopold C, Schandry R. 2001. The heartbeat-evoked brain potential in patients suffering from diabetic neuropathy and in healthy control persons. *Clin. Neurophysiol.* 112(4):674–82

Liang F, Paulo R, Molina G, Clyde MA, Berger JO. 2008. Mixtures of g-priors for Bayesian variable selection. *J. Am. Stat. Assoc.* 103(481):410–23

Longo MR, Schüür F, Kammers MPM, Tsakiris M, Haggard P. 2008. What is embodiment? A psychometric approach. *Cognition.* 107(3):978–98

Luft CDB, Bhattacharya J. 2015. Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. *Sci. Rep.* 5:15717

MacKinnon S, Gevirtz R, McCraty R, Brown M. 2013. Utilizing heartbeat evoked potentials to identify cardiac regulation of vagal afferents during emotion and resonant breathing. *Appl. Psychophysiol. Biofeedback.* 38(4):241–55

Macuga KL, Frey SH. 2012. Neural representations involved in observed, imagined, and

imitated actions are dissociable and hierarchically organized. *Neuroimage*. 59(3):2798–2807

Maguire EA, Burgess N, Donnett JG, Frackowiak RSJ, Frith CD, Keefe JO. 1998. Knowing where and getting there : A human navigation network. *Science*. 280:921–24

Makin TR, Holmes NP, Ehrsson HH. 2008. On the other hand: Dummy hands and peripersonal space. *Behav. Brain Res*. 191(1):1–10

Margulies DS, Vincent JL, Kelly C, Lohmann G, Uddin LQ, et al. 2009. Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc Natl Acad Sci USA*. 106(47):20069–74

Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods*. 164(1):177–90

Martinelli P, Sperduti M, Piolino P. 2013. Neural substrates of the self-memory system: New insights from a meta-analysis. *Hum. Brain Mapp*. 34(7):1515–29

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN. 2007. Wandering minds: the default network and stimulus-independent thought. *Science*. 315(5810):393–95

Matthias E, Schandry R, Duschek S, Pollatos O. 2009. On the relationship between interoceptive awareness and the attentional processing of visual stimuli. *Int. J. Psychophysiol.* 72(2):154–59

Mazzola L, Lopez C, Faillenot I, Chouchou F, Mauguière F, Isnard J. 2014. Vestibular responses to direct stimulation of the human insular cortex. *Ann. Neurol.* 76(4):609–19

Medford N, Critchley HD. 2010. Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct. Funct*. 214(5–6):535–49

Miele DB, Wager TD, Mitchell JP, Metcalfe J. 2011. Dissociating neural correlates of action monitoring and metacognition of agency. *J. Cogn. Neurosci.* 23(11):3620–36

Montoya P, Schandry R, Müller A. 1993. Heartbeat evoked potentials (HEP): topography and influence of cardiac awareness and focus of attention. *Electroencephalogr. Clin. Neurophysiol.* 88(3):163–72

Moore JW, Ruge D, Wenke D, Rothwell J, Haggard P. 2010. Disrupting the experience of control in the human brain: pre-supplementary motor area contributes to the sense of agency. *Proc. R. Soc. B*. 277:2503–9

Murray RJ, Schaer M, Debbané M. 2012. Degrees of separation: a quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self-

and other-reflection. *Neurosci. Biobehav. Rev.* 36(3):1043–59

Nagai Y, Critchley HD, Featherstone E, Trimble MR, Dolan RJ. 2004. Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: a physiological account of a "default mode" of brain function. *Neuroimage.* 22(1):243–51

Naito E, Morita T, Amemiya K. 2016. Body representations in the human brain revealed by kinesthetic illusions and their essential contributions to motor control and corporeal awareness. *Neurosci. Res.* 104:16–30

Neubert F-X, Mars RB, Sallet J, Rushworth MFS. 2015. Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 1–10

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. *Neuroimage.* 31(1):440–57

Nosaka S, Murase S, Murata K, Inui K. 1995. "Aortic baroreceptor" neurons in the nucleus tractus solitarius in rats: convergence of cardiovascular inputs as revealed by heartbeat-locked activity. *J. Auton. Nerv. Syst.* 55(1–2):69–80

O'Brien WH, Reid GJ, Jones KR. 1998. Differences in heartbeat awareness among males with higher and lower levels of systolic blood pressure. *Int. J. Psychophysiol.* 29(1):53–63

Ochsner KN, Knierim K, Ludlow DH, Hanelin J, Ramachandran T, et al. 2004. Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J. Cogn. Neurosci.* 16(10):1746–72

Ohl S, Wohltat C, Kliegl R, Pollatos O, Engbert R. 2016. Microsaccades are coupled to heartbeat. *J. Neurosci.* 36(4):1237–41

Oostenveld R, Fries P, Maris E, Schoffelen JM. 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011(156869):

Park H-D, Bernasconi F, Bello-Ruiz J, Pfeiffer C, Salomon R, Blanke O. 2016. Transient modulations of neural responses to heartbeats covary with bodily self-consciousness. *J. Neurosci.* 36(32):8453–60

Park H-D, Correia S, Ducorps A, Tallon-Baudry C. 2014. Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nat. Neurosci.* 17(4):612–18

Park H-D, Tallon-Baudry C. 2014. The neural subjective frame: from bodily signals to perceptual consciousness. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369(1641):20130208

Park S, Won MJ, Lee EC, Mun S, Park M-C, Whang M. 2015. Evaluation of 3D cognitive fatigue using heart-brain synchronization. *Int. J. Psychophysiol.* 97(2):120–30

Perrin F, Maquet P, Peigneux P, Ruby P, Degueldre C, et al. 2005. Neural mechanisms involved in the detection of our first name: A combined ERPs and PET study. *Neuropsychologia*. 43(1):12–19

Perrin F, Schnakers C, Schabus M, Degueldre C, Goldman S, et al. 2006. Brain response to one's own name in vegetative state, minimally conscious state, and locked-in syndrome. *Arch. Neurol.* 63(4):562–69

Petkova VI, Björnsdotter M, Gentile G, Jonsson T, Li TQ, Ehrsson HH. 2011. From part- to whole-body ownership in the multisensory brain. *Curr. Biol.* 21(13):1118–22

Pfeiffer C, Lopez C, Schmutz V, Duenas JA, Martuzzi R, Blanke O. 2013. Multisensory Origin of the Subjective First-Person Perspective: Visual, Tactile, and Vestibular Mechanisms. *PLoS One*. 8(4):e61751

Philippi CL, Tranel D, Duff M, Rudrauf D. 2015. Damage to the default mode network disrupts autobiographical memory retrieval. *Soc. Cogn. Affect. Neurosci.* 10(3):318–26

Phillips GC, Jones GE, Rieger EJ, Snell JB. 1999. Effects of the presentation of false heart-rate feedback on the performance of two common heartbeat-detection tasks. *Psychophysiology*. 36(4):504–10

Platek SM, Loughead JW, Gur RC, Busch S, Ruparel K, et al. 2006. Neural substrates for functionally discriminating self-face from personally familiar faces. *Hum. Brain Mapp.* 27(2):91–98

Plotnik JM, de Waal FBM, Reiss D. 2006. Self-recognition in an Asian elephant. *Proc. Natl. Acad. Sci. U. S. A.* 103(45):17053–57

Pollatos O, Herbert BM, Mai S, Kammer T. 2016. Changes in interoceptive processes following brain stimulation. *Philos. Trans. R. Soc. B.* 371(1708):1–11

Pollatos O, Kirsch W, Schandry R. 2005a. Brain structures involved in interoceptive awareness and cardioafferent signal processing: a dipole source localization study. *Hum. Brain Mapp.* 26(1):54–64

Pollatos O, Kirsch W, Schandry R. 2005b. On the relationship between interoceptive awareness, emotional experience, and brain processes. *Cogn. brain Res.* 25(3):948–62

Pollatos O, Kurz AL, Albrecht J, Schreder T, Kleemann AM, et al. 2008. Reduced perception of bodily signals in anorexia nervosa. *Eat. Behav.* 9(4):381–88

Pollatos O, Matthias E, Schandry R. 2007a. Heartbeat perception and P300 amplitude in a

visual oddball paradigm. *Clin. Neurophysiol.* 118(10):2248–53

Pollatos O, Schandry R. 2004. Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. *Psychophysiology*. 41(3):476–82

Pollatos O, Schandry R, Auer DP, Kaufmann C. 2007b. Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Res.* 1141:178–87

Qin P, Northoff G. 2011. How is our self related to midline regions and the default-mode network? *Neuroimage*. 57(3):1221–33

Reiss D, Marino L. 2001. Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proc. Natl. Acad. Sci.* 98(10):5937–42

Richter CG, Babo-Rebelo M, Schwartz D, Tallon-Baudry C. 2017. Phase-amplitude coupling at the organism level: The amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm. *Neuroimage*. 146:951–58

Richter FR, Cooper RA, Bays PM, Simons JS. 2016. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *Elife*. 5:e18260

Rochat P. 2003. Five levels of self-awareness as they unfold early in life. *Conscious. Cogn.* 12(4):717–31

Ronchi R, Bello-Ruiz J, Lukowska M, Herbelin B, Cabrilo I, et al. 2015. Right insular damage decreases heartbeat awareness and alters cardio-visual effects on bodily self-consciousness. *Neuropsychologia*. 70:11–20

Ruby P, Decety J. 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4(5):546–50

Saari M, Pappas B. 1976. Cardiac cycle phase and movement and reaction times. *Percept Mot Ski.* 42(3):767–70

Salomon R, Ronchi R, Donz J, Bello-Ruiz J, Herbelin B, et al. 2016. The insula mediates access to awareness of visual stimuli presented synchronously to the heartbeat. *J. Neurosci.* 36(18):5115–27

Saper CB. 2002. The central autonomic nervous system: conscious visceral perception and autonomic pattern generation. *Annu. Rev. Neurosci.* 25:433–69

Sartre J-P. 1943. *L'être et Le Néant.* Gallimard ed.

Schandry R. 1981. Heart beat perception and emotional experience. *Psychophysiology*. 18(4):483–88

Schandry R, Montoya P. 1996. Event-related brain potentials and the processing of cardiac

activity. *Biol. Psychol.* 42(1–2):75–85

Schandry R, Sparrer B, Weitkunat R. 1986. From the heart to the brain: a study of heartbeat contingent scalp potentials. *Int. J. Neurosci.* 30(4):261–75

Schandry R, Weitkunat R. 1990. Enhancement of heartbeat-related brain potentials through cardiac awareness training. *Int. J. Neurosci.* 53(2–4):243–53

Schmitz TW, Kawahara-Baccus TN, Johnson SC. 2004. Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *Neuroimage.* 22(2):941–47

Schneider F, Bermpohl F, Heinzel A, Rotte M, Walter M, et al. 2008. The resting brain and our self: self-relatedness modulates resting state neural activity in cortical midline structures. *Neuroscience.* 157(1):120–31

Schooler JW, Smallwood J, Christoff K, Handy TC, Reichle ED, Sayette MA. 2011. Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cogn. Sci.* 15(7):319–26

Schulz A, Ferreira de Sá DS, Dierolf AM, Lutz A, van Dyck Z, et al. 2015a. Short-term food deprivation increases amplitudes of heartbeat-evoked potentials. *Psychophysiology.* 52(5):695–703

Schulz A, Köster S, Beutel ME, Schächinger H, Vögele C, et al. 2015b. Altered patterns of heartbeat-evoked potentials in depersonalization/derealization disorder. *Psychosom. Med.* 77(5):506–16

Schulz A, Lass-Hennemann J, Sütterlin S, Schächinger H, Vögele C. 2013a. Cold pressor stress induces opposite effects on cardioceptive accuracy dependent on assessment paradigm. *Biol. Psychol.* 93(1):167–74

Schulz A, Strelzyk F, Ferreira de Sá DS, Naumann E, Vögele C, Schächinger H. 2013b. Cortisol rapidly affects amplitudes of heartbeat-evoked brain potentials-Implications for the contribution of stress to an altered perception of physical sensations? *Psychoneuroendocrinology.* 38(11):2686–93

Schulz SM. Neural Correlates of Heart-Focussed Interoception: an fMRI Meta-Analysis. *Philos. Trans. R. Soc. B*

Sel A, Azevedo RT, Tsakiris M. 2016. Heartfelt self: cardio-visual integration affects self-face recognition and interoceptive cortical processing. *Cereb. Cortex.* 1–12

Sforza A, Bufalari I, Haggard P, Aglioti SM. 2010. My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Soc. Neurosci.* 5(2):148–62

Shah P, Hall R, Catmur C, Bird G. 2016. Alexithymia, not autism, is associated with impaired

interoception . *Cortex*. 81:215–20

Shao S, Shen K, Wilder-Smith EP V, Li X. 2011. Effect of pain perception on the heartbeat evoked potential. *Clin. Neurophysiol.* 122(9):1838–45

Shepherd JT. 1985. The heart as sensory organ. *J. Am. Coll. Cardiol.* 5(6):83B–87B.

Shmueli K, van Gelderen P, de Zwart JA, Horovitz SG, Fukunaga M, et al. 2007. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. *Neuroimage*. 38(2):306–20

Simmons WK, Avery JA, Barcalow JC, Bodurka J, Drevets WC, Bellgowan P. 2013. Keeping the body in mind: insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Hum. Brain Mapp.* 34(11):2944–58

Smallwood J, Schooler JW. 2015. The science of mind wandering: empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* 66(1):487–518

Sokoloff L, Mangold R, Wechsler RL, Kenney C, Kety SS. 1955. The effect of mental arithmetic on cerebral circulation and metabolism. *J. Clin. Invest.* 34(7, Part 1):1101–8

Sperduti M, Delaveau P, Fossati P, Nadel J. 2011. Different brain structures related to self- and external-agency attribution: A brief review and meta-analysis. *Brain Struct. Funct.* 216:151–57

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. 1983. *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press

Spreng RN, Mar RA, Kim ASN. 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21(3):489–510

Stawarczyk D, Majerus S, Van der Linden M, D'Argembeau A. 2012. Using the Daydreaming Frequency Scale to Investigate the Relationships between Mind-Wandering, Psychological Well-Being, and Present-Moment Awareness. *Front. Psychol.* 3(September):363

Strigo IA, Craig AD. 2016. Interoception, homeostatic emotions and sympathovagal balance. *Philos. Trans. R. Soc. London B Biol. Sci.* 371(1708):1–9

Suarez SD, Gallup GG. 1981. Self-recognition in chimpanzees and orangutans, but not gorillas. *J. Hum. Evol.* 10(2):175–88

Suddendorf T, Butler DL. 2013. The nature of visual self-recognition. *Trends Cogn. Sci.* 17(3):121–27

Sugiura M, Sassa Y, Jeong H, Miura N, Akitsuki Y, et al. 2006. Multiple brain networks for visual self-recognition with different sensitivity for motion and body part. *Neuroimage*. 32(4):1905–17

Summerfield JJ, Hassabis D, Maguire EA. 2009. Cortical midline involvement in autobiographical memory. *Neuroimage*. 44(3):1188–1200

Suzuki K, Garfinkel SN, Critchley HD, Seth AK. 2013. Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia*. 51(13):2909–17

Svoboda E, McKinnon MC, Levine B. 2006. The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*. 44(12):2189–2208

Synofzik M, Vosgerau G, Newen A. 2008. Beyond the comparator model: A multifactorial two-step account of agency. *Conscious. Cogn.* 17:219–39

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM. 2011. Brainstorm: A user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011:879716

Terhaar J, Viola FC, Bär KJ, Debener S. 2012. Heartbeat evoked potentials mirror altered body perception in depressed patients. *Clin. Neurophysiol.* 123(10):1950–57

Teves D, Videen TO, Cryer PE, Powers WJ. 2004. Activation of human medial prefrontal cortex during autonomic responses to hypoglycemia. *Proc. Natl. Acad. Sci. U. S. A.* 101(16):6217–21

Thayer JF, Ahs F, Fredrikson M, Sollers JJ, Wager TD. 2012. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neurosci. Biobehav. Rev.* 36(2):747–56

Tian L, Jiang T, Liu Y, Yu C, Wang K, et al. 2007. The relationship within and between the extrinsic and intrinsic systems indicated by resting state correlational patterns of sensory cortices. *Neuroimage*. 36(3):684–90

Trapnell PD, Campbell JD. 1999. Private Self-Consciousness and the Five-Factor Model of Personality: Distinguishing Rumination From Reflection. *J. Pers. Soc. Psychol.* 76(2):284–304

Tsakiris M. 2008. Looking for myself: Current multisensory input alters self-face recognition. *PLoS One*. 3(12):e4040

Tsakiris M. 2010. My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia*. 48(3):703–12

Tsakiris M, Haggard P. 2005. The rubber hand illusion revisited: visuotactile integration and

self-attribution. *J Exp Psychol Hum Percept Perform*. 31(1):80–91

Tsakiris M, Hesse MD, Boy C, Haggard P, Fink GR. 2007a. Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb. cortex*. 17(10):2235–44

Tsakiris M, Schütz-Bosbach S, Gallagher S. 2007b. On agency and body-ownership: phenomenological and neurocognitive reflections. *Conscious. Cogn.* 16(3):645–60

Tsakiris M, Tajadura-Jiménez A, Costantini M. 2011. Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. *Proc. Biol. Sci.* 278(1717):2470–76

Turk DJ, Heatherton TF, Kelley WM, Funnell MG, Gazzaniga MS, Macrae CN. 2002. Mike or me? Self-recognition in a split-brain patient. *Nat. Neurosci.* 5(9):841–42

Tusche A, Smallwood J, Bernhardt BC, Singer T. 2014. Classifying the wandering mind: revealing the affective content of thoughts during task-free rest periods. *Neuroimage*. 97:107–16

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, et al. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 15(1):273–89

Uddin LQ, Molnar-Szakacs I, Zaidel E, Iacoboni M. 2006. rTMS to the right inferior parietal lobule disrupts self-other discrimination. *Soc Cogn Affect Neurosci.* 1(1):65–71

Ueno A, Hirata S, Fuwa KK, Sugama K, Kusunoki K, et al. 2010. Brain activity in an awake chimpanzee in response to the sound of her own name. *Biol. Lett.* 6(3):311–13

Umeda S, Tochizawa S, Shibata M, Terasawa Y. 2016. Prospective memory mediated by interoceptive accuracy: a psychophysiological approach. *Philos. Trans. R. Soc. B.* 371(1708):1–8

Vaitl D. 1996. Interoception. *Biol. Psychol.* 42(1–2):1–27

van der Meer L, Costafreda S, Aleman A, David AS. 2010. Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neurosci. Biobehav. Rev.* 34(6):935–46

Vanhaudenhuyse A, Demertzi A, Schabus M, Noirhomme Q, Bredart S, et al. 2011. Two distinct neuronal networks mediate the awareness of environment and of self. *J. Cogn. Neurosci.* 23(3):570–78

Vann SD, Aggleton JP, Maguire EA. 2009. What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* 10(11):792–802

Vianna EPM, Naqvi N, Bechara A, Tranel D. 2009. Does vivid emotional imagery depend on body signals? *Int. J. Psychophysiol.* 72(1):46–50

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CMA. 2010. The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. *Neuroimage.* 51(1):112–22

Vogeley K, Fink GR. 2003. Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* 7(1):38–42

Vogeley K, May M, Ritzl A, Falkai P, Zilles K, Fink GR. 2004. Neural correlates of first-person perspective as one constituent of human self-consciousness. *J. Cogn. Neurosci.* 16(5):817–27

Wada M, Takano K, Ora H, Ide M, Kansaku K. 2016. The rubber tail illusion as evidence of body ownership in mice. *J. Neurosci.* 36(43):11133–37

Weiss C, Tsakiris M, Haggard P, Schütz-Bosbach S. 2014. Agency in the sensorimotor system and its relation to explicit action awareness. *Neuropsychologia*. 52:82–92

Werner NS, Jung K, Duschek S, Schandry R. 2009. Enhanced cardiac perception is associated with benefits in decision-making. *Psychophysiology*. 46(6):1123–29

Wetzels R, Wagenmakers E-J. 2012. A default Bayesian hypothesis test for correlations and partial correlations. *Psychon. Bull. Rev*. 19:1057–64

Whitehead W, Drescher V. 1980. Perception of gastric contractions and self-control of gastric motility. *Psychophysiology*. 17(6):552–58

Whitehead W, Drescher V, Heiman P, Blackwell B. 1977. Relation of heart rate control to heartbeat perception. *Biofeedback Self Regul.* 2(4):317–92

Whitfield-Gabrieli S, Moran JM, Nieto-Castañón A, Triantafyllou C, Saxe R, Gabrieli JDE. 2011. Associations and dissociations between default and self-reference networks in the human brain. *Neuroimage*. 55:225–32

Wiebking C, Duncan NW, Tiret B, Hayes DJ, Marjańska M, et al. 2014. GABA in the insula - a predictor of the neural response to interoceptive awareness. *Neuroimage*. 86:10–18

Wilson BA, Berry E, Gracey F, Harrison C, Stow I, et al. 2005. Egocentric disorientation following bilateral parietal lobe damage. *Cortex*. 41(4):547–54

Wood N, Cowan N. 1995. The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *J. Exp. Psychol.* 21(1):255–60

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated

synthesis of human functional neuroimaging data. *Nat. Methods*. 8(8):665–70

Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, et al. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106(3):1125–65

Yuan H, Zotev V, Phillips R, Bodurka J. 2013. Correlated slow fluctuations in respiration, EEG, and BOLD fMRI. *Neuroimage*. 79:81–93

Zahavi D. 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge. The MIT Pr ed.

Zaki J, Davis JI, Ochsner KN. 2012. Overlapping activity in anterior insula during interoception and emotional experience. *Neuroimage*. 62(1):493–99

Zentgraf K, Stark R, Reiser M, Künzell S, Schienle A, et al. 2005. Differential activation of pre-SMA and SMA proper during action observation: Effects of instructions. *Neuroimage*. 26(3):662–72