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

cognitive role in selfhood and a physiological role in autonomic function, although sometimes noted (Buckner et al., 2008; Iacovella 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. 1B). 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. 1C), 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. 1A). 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: $1.80 \pm 0.032$ 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 = $1.8 \pm 0.5\%$, "Me" scale = $1.2 \pm 0.4\%$, Time = $8.0 \pm 2.0\%$, Valence = $5.1 \pm 1.3\%$). 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, $\beta$ 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 $\beta$ 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 Brain-Storm 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. 1B). 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 |
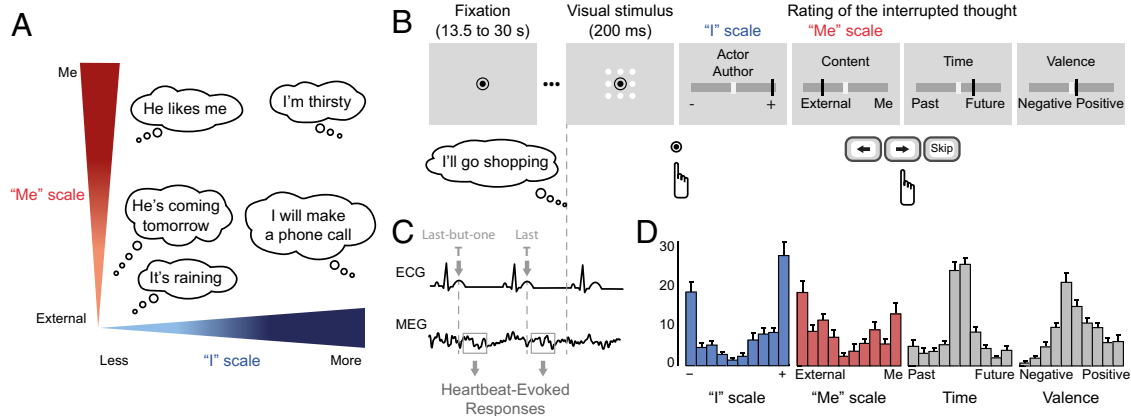


**Figure 1.** Experimental paradigm and behavior. **A**, Examples of thoughts along the two scales of self-relatedness. The "Me" scale described the content of the thought oriented either toward oneself or toward an external object, event, or person. The "I" scale described the engagement of the participant 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 1D. 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 $r = 0.85 ± 0.06$, 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 $r = 0.14 ± 0.03$, $t_{(15)} = 5.21$, 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. 2A,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. 2C; "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
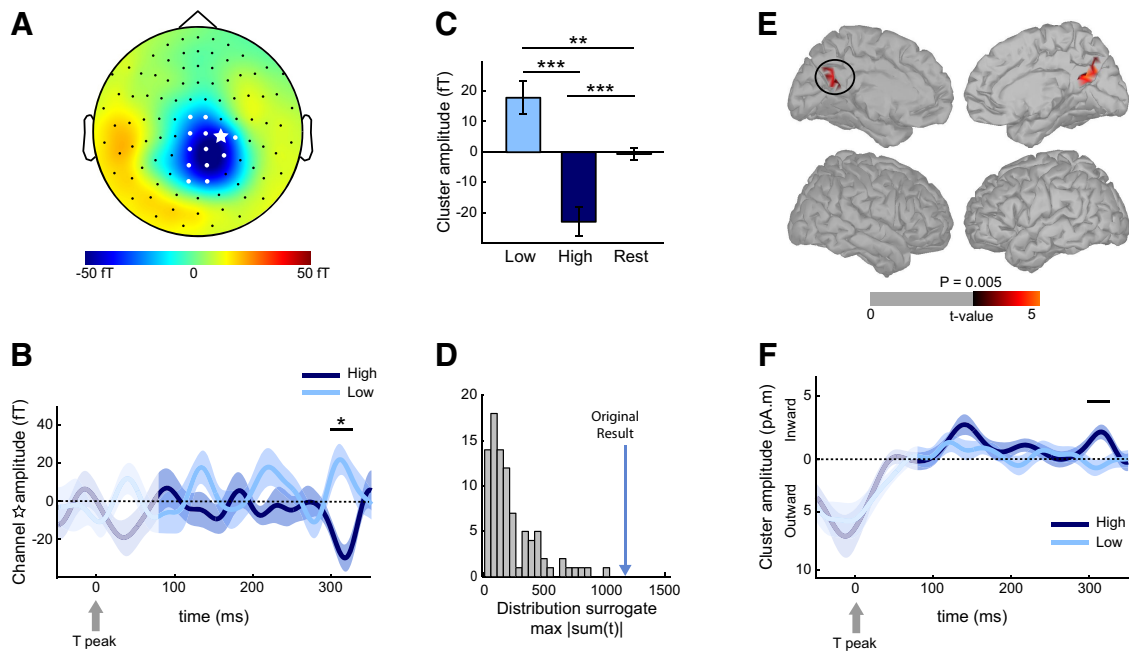
**Figure 2.** Differential HERs for "high" and "low" ratings on the "I" scale. ***A***, Topographical map of the HER difference between "high" and "low" ratings on the "I" scale, grand-averaged across 16 participants, in the 298 –327 ms time window in which a significant difference was observed (Monte Carlo $p = 0.0313$, corrected for multiple comparisons). White dots represent the sensors contributing to the significant cluster. ***B***, Time course of the HER ($\pm$ 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 ($\pm$ 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": $9.3 \pm 1.9$ fT, baseline: $3.4 \pm 1.2$ fT, "low": $-5.7 \pm 2.1$ 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 $\leq 3.52$), where the regressors were orthogonalized to separate shared from unique variance. The $\beta$ 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 $\beta$ 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.
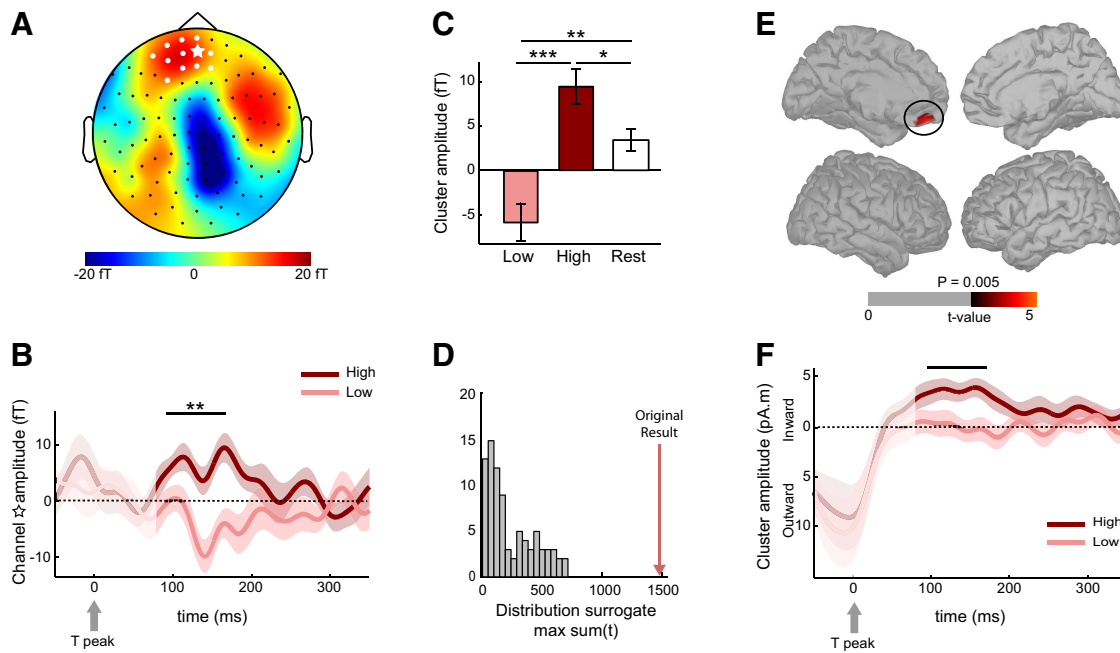
**Figure 3.** Differential HERs for "high" and "low" ratings on the "Me" scale. *A*, Topographical map of the HER difference between "high" and "low" ratings on the "Me" scale, grand-averaged across 16 participants, in the 94–169 ms time window in which a significant difference was observed (Monte Carlo 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 × 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.
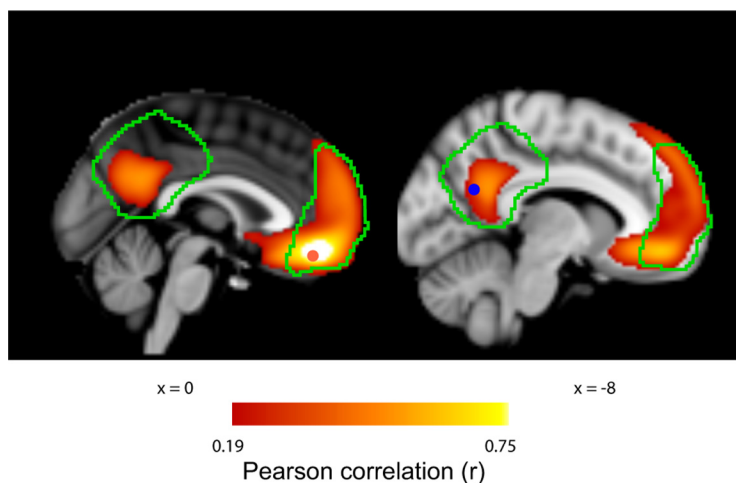


**Figure 4.** Functional connectivity between vmPFC and vPC and overlap with default network (DN). Red-white represents functional connectivity computed from resting-state BOLD time series of 1000 subjects at rest (Yarkoni et al., 2011), with a seed placed in left vmPFC (MNI coordinates: 0, 45, −15, 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. 2E,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 cm²). 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.) | $-17 \pm 31$ | $-19 \pm 40$ | $t_{(15)} = 0.028$ $p = 0.98$ | Substantial (4.40) | $-6.7 \pm 22$ | $-30 \pm 35$ | $t_{(15)} = 0.44$ $p = 0.67$ | 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": $-2.03 \pm 0.50$ 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": $0.57 \pm 0.52$ 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": $-3.69 \pm 0.72$ pA.m, $t_{(15)} = -5.32$, $p = 1.7 \times 10^{-4}$, "low": $-0.040 \pm 0.86$ 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 $\geq 3.58$, indicating substantial evidence for the ab-

sence 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 $\geq 3.68$), pupil diameter (Fig. 5B; both $p \geq 0.67$, Bayes factors $\geq 3.78$), alpha power (Fig. 5C; 8–12 Hz, averaged over occipitoparietal sensors, both $p \geq 0.41$, Bayes factors $\geq 2.57$). Last, the number of blinks (both $p \geq 0.75$, Bayes factor $\geq 4.05$) and saccades (both $p \geq 0.38$, Bayes factors $\geq 2.40$) 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. 2D, 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. 5A). 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
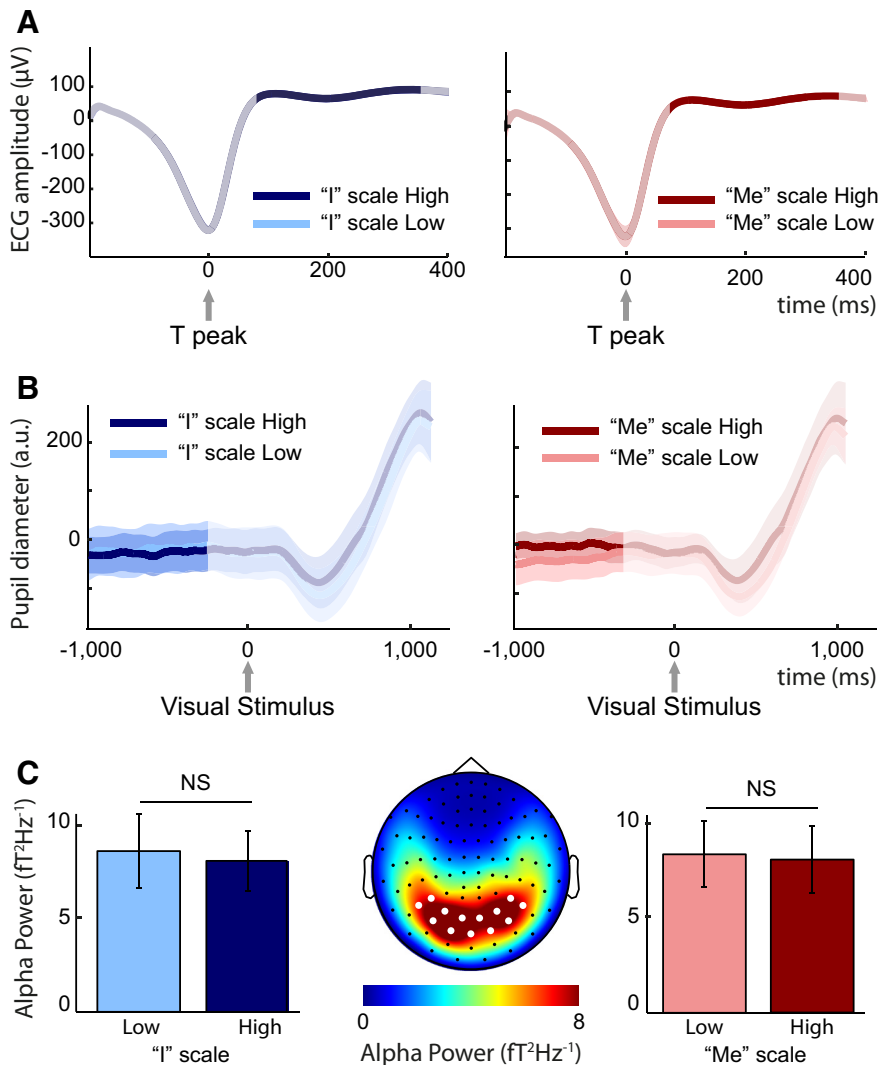
**Figure 5.** Controls. ***A***, Time course of the average ECG signal (± SEM) for "high" and "low" ratings along the "I" (left) and "Me" (right) scales, on the vertical derivation lead II. The signal appearing in darker color corresponds to the time window that was analyzed in the MEG data. The ECG, 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]**

| | Scores (mean ± SEM) | "I" cluster | | "Me" cluster | |
|---|---|---|---|---|---|
| | | 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