



Linking bodily, environmental and mental states in the self—A three-level model based on a meta-analysis

Pengmin Qin^{a,b,*}, Mingxia Wang^{c,1}, Georg Northoff^{d,e}

^a Centre for Studies of Psychological Applications, Guangdong Key Laboratory of Mental Health and Cognitive Science, School of Psychology, South China Normal University, Guangzhou, Guangdong, China

^b Pazhou Lab, Guangzhou, 510335, China

^c Hearing and Speech Department, Xinhua College of Sun Yat-sen University, Guangzhou, China

^d Institute of Mental Health Research, University of Ottawa, Ottawa, Ontario, Canada

^e Mental Health Centre, Zhejiang University School of Medicine, Hangzhou, China

ARTICLE INFO

Keywords:

Insula
TPJ
AMPFC
Familiarity
Interoception
Exteroception

ABSTRACT

Current researchers mostly agree that the self consists of both bodily and non-bodily environmental information. The neural mechanism underlying the integration of this information remains unclear. In this study, we propose a neural model subdividing self-processing into three intimately connected levels with different extension: Interoceptive-processing, Exteroceptive-processing and Mental-self-processing. We applied ALE meta-analyses on neuroimaging studies to analyze their neural patterns. Our results show common involvement of insula across all three levels including differentiation of self and familiarity. Common activities in Exteroceptive- and Mental-self-processing were found in the anteromedial prefrontal cortex (AMPFC) and the temporal parietal junction (TPJ), suggesting that the two regions likely serve basic functions in differentiation and integration of self-other information. Finally, Mental-self-processing involves extensive regions such as the cingulate cortex and medial prefrontal cortex, in addition to the insula, AMPFC and TPJ, which could specialize in adding self-relatedness to environment information. We conclude that there is a gradient organization in self-processing, through which body-environment information is integrated for the self via propagation from Interoceptive-processing to Mental-self-processing.

1. Introduction

What is the self? The self has long been an important concept, debated by philosophers for centuries. In modern psychology, William James argued that the self is the fundamental unit of analysis for a science of mental life, the problem about which everything else revolves (James, 1890). With its obvious importance, efforts spent in investigating the self in psychology, neuroscience and neuropsychiatry have grown rapidly in recent years. In the last three decades, various studies about interoception (Park et al., 2016), body-related stimuli (e.g. one's own face or one's own agency of an action) (van Veluw and Chance, 2014), and self-related abstract stimuli/external environment stimuli (e.g. trait adjectives, geometrical figure) (Hu et al., 2016) lead to the idea that the self may intrinsically combine the body and natural/social environment (Craig, 2010; Park and Blanke, 2019; Qin and Northoff, 2011). For instance, it is proposed that pre-reflective selfhood

emerges from one's everyday experience through his/her body (Apps and Tsakiris, 2014; Limanowski and Blankenburg, 2013), from which external stimuli relevant to the self could be integrated and processed with better performance through processes like self-referential/related processing (de Greck et al., 2008; Northoff, 2016a, 2011, 2007; Sui and Humphreys, 2016, 2015). Considering the importance of the integration between bodily and external environment information for the self, we here propose a neural model of self which is based on a mechanism of how external stimuli can become self-related and thereby integrated within the self. While, at the same time, the self becomes thereby extended beyond its body to the external environment.

Upon reviewing the literature with a specific focus on the kind of paradigms used to elicit and investigate the self, we identified three levels of neural processing, through which self-related information from the internal body to the external environment is propagated and integrated within the self. In the following paragraphs, each proposed

* Corresponding author at: Centre for Studies of Psychological Applications, Guangdong Key Laboratory of Mental Health and Cognitive Science, School of Psychology, South China Normal University, Guangzhou, Guangdong, China.

E-mail address: qin.pengmin@m.scun.edu.cn (P. Qin).

¹ These authors contribute equally as first authors to this paper.

<https://doi.org/10.1016/j.neubiorev.2020.05.004>

Received 3 August 2019; Received in revised form 6 April 2020; Accepted 8 May 2020

Available online 31 May 2020

0149-7634/ © 2020 Elsevier Ltd. All rights reserved.

Three-level Model of the Self

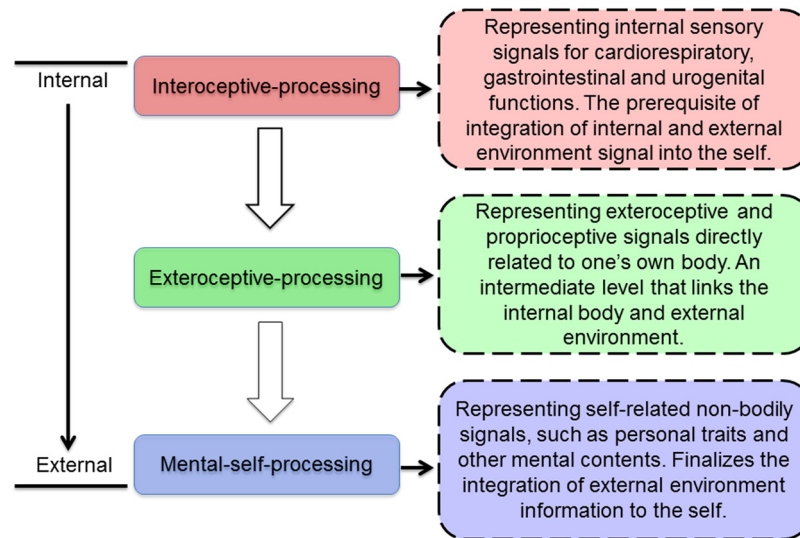


Fig. 1. A schematic representation of the three-level-processing model of self, summarizing the primary role of each level in self-processing, as well as the contents of the neural representations involved in each level.

level of self-processing will be interpreted in an internal to external sequence.

- 1) **Interoceptive-processing.** Several previous theories proposed that the self is grounded on interoceptive processing, such as the neural representation of cardiorespiratory, gastrointestinal and urogenital signals (Craig, 2009; Park et al., 2014; Seth, 2013; Tsakiris and Critchley, 2016). These theories have been supported by increasing empirical evidence. For instance, it has been found that heartbeat-evoked-responses recorded with MEG covaried with self-related spontaneous thoughts (Babo-rebelo et al., 2016), as well as bodily self-consciousness induced by the full-body illusion tasks (Park et al., 2016). Other studies have also found that the experience of body ownership and self-identification could be modulated by interoceptive signals (Sel et al., 2017; Seth, 2013; Suzuki et al., 2013). Furthermore, the fact that some of these interoceptive modulations do not involve attention, indicates that interoception might affect self-processing in an implicit and automatic way. These findings implicated a causal role of interoception in self-awareness by highlighting its importance in the processing of external self-related information. Based on the above theories and evidence, in the current study, we propose that interoceptive-processing is one level of bodily-self-processing, and is the prerequisite of the integration of bodily and external environment information for self-processing.
- 2) **Exteroceptive-processing.** Having to interact with the outer world, our body cannot rely on internal senses alone, i.e., its interoceptive stimuli. As reflected in the “mirror tests”, the ability in self-recognition is one of the crucial markers for development of self-awareness (Anderson, 1984). It involves processing external stimuli which were the direct projections of one’s own bodily signals in the external environment, such as photo of one’s own face and other body parts, or a cursor/mouse movement driven by one’s own action. Such processing directly links the body with the external environment, which is based on various exteroceptive signals such as vision and touch, as well as multisensory signals driven directly by our own physical actions, such as the combination of synchronous visual feedback and proprioceptive signals in an agency task (Sperduti et al., 2011). Evidence showed that exteroceptive signals (including proprioceptive signals) about our own body, as well as integration of interoceptive and exteroceptive signals, are crucial to modulating basic self-other boundaries (Park and Blanke, 2019);

that, in turn, seems to constitute the very basis of social interactions (Sperduti et al., 2011; Suzuki et al., 2013; Tsakiris, 2017). In the current study, we define this level of processing as exteroceptive-processing, and we propose that it incorporates exteroceptive, proprioceptive, and multisensory signals for the self, and links the intero- and exteroceptive body with external environment information which show a direct relationship with the body as mentioned above.

- 3) **Mental-self-processing.** This level of processing extends the contents of the self to external stimuli which do not have any direct contact with the body, such as one’s own name or cell phone. These more mental and virtual (rather than body-based physical) stimuli could encompass infinite varieties of information, such as self-related traits, one’s own name, autobiographical memory, first-person perspective judgments, etc. This level of processing has been frequently implicated in the so-called self-referential effect, in which self-related information, such as trait words referred to oneself, could be processed with better performance, such as better memory (Kim and Johnson, 2014; Rogers et al., 1977; Shi et al., 2011; Sui et al., 2013). However, it has long been proposed that the processing of non-bodily signals such as in social cognition, could be influenced by bodily signals (Farmer et al., 2014; Maister et al., 2013). The above evidence indicates that self-processing of external non-bodily signals, even abstract ones, could be grounded on bodily signals. In the current study, we define this level of processing as Mental-self-processing. We propose that this level of processing represents the abstract affiliation of the external environment with the self in mental contents (i.e. self-relatedness of the external environment stimuli), by incorporating Interoceptive-processing, Exteroceptive-processing, and finally realizes the integration of bodily and external environment information which show non-direct contact with body as mentioned above.

The proposed functions and neural contents for each level of self-processing in our model are summarized in Fig. 1. In the following paragraphs, we will outline the experimental designs involved in each of these processing levels.

1.1. Experiments on interoceptive-processing

Interoception processes internal sensory signals such as thirst, itch, heartbeat, distension of the bladder, stomach, etc. (Craig, 2009).

However, given the difficulty in manipulating a person's internal physical states, currently, only a few of these fields have been sufficiently investigated to produce relatively reliable findings. With reference to several review studies on interoception (Brener and Ring, 2016; Craig, 2009; Tsakiris and Critchley, 2016), we've selected the major physiological functions as representatives for interoceptive-processing: cardiorespiratory, gastrointestinal and urogenital functions. Specific paradigms investigating the neural mechanisms of these functions include: heartbeat detection/differentiation, dyspnea/hyperpnea, hunger/thirst, gastric distension and micturition. These paradigms generally involve some form of manipulation to elicit changes of the corresponding physiological states so that neural activities in the different states could be compared. In the following paragraphs, more details of the different paradigms are introduced:

For cardiorespiratory functions, heartbeat detection/differentiation is the most commonly adopted paradigm. This approach usually requires the participants to differentiate their own heartbeat from other tones, which investigates interoceptive accuracy (Stern et al., 2017), or analyzes the participants' heart rate change driven by task-induced stress (e.g. speech preparation), investigating interoceptive sensibility (Garfinkel et al., 2016; Suzuki et al., 2013). As compared to heartbeat, neural mechanisms of respiration were far less investigated in especially neuroimaging studies, due to its major contribution in creating BOLD signals artifacts (Murphy et al., 2013). Two major paradigms currently used are dyspnea (air hunger/shortness of breath) and less commonly, hyperpnea (hyperventilation). For example, in one study, various levels of breathing difficulty were induced via an MRI-compatible breathing circuit that presents resistive loads to the inspiratory end (Stoekel et al., 2015).

For gastrointestinal functions, sensory signals conveyed by the stomach and the intestines contribute significantly to driving eating behaviors (Stevenson et al., 2015). Hunger and thirst are the two major interoceptive sensory signals that serve such motivational function. In hunger studies, participants were often required to differentiate between food and non-food related stimuli (visual, taste or olfactory) after various length of fasting (Frank et al., 2010); a lesser used approach involves comparing brain activities between fasted states and satiated states (Haase et al., 2009). In thirst studies, participants were usually injected with hypertonic saline to induce different levels of thirst, and were monitored for their neural change at rest or during task (discriminating beverage/non-beverage stimuli) (Egan et al., 2003). Apart from hunger and thirst, distention from digestive organs also provides an important source of sensory signals that contribute significantly to satiation (Wang et al., 2008). This paradigm involves distending the gastric volume using experimental manipulations, usually via a gastric balloon (Wang et al., 2008) or sometimes infusion of water or other nutrients (Camps et al., 2018).

For urogenital functions, we primarily selected the micturition paradigm. In this paradigm, sensations from the bladder are studied via manipulating the bladder capacity by naturally (Kuhntz-Buschbeck et al., 2009) or manually (Mehnert et al., 2011) filling the bladder to induce different levels of desires to void. Please note that other than micturition, genital stimulation/sexual arousal are also important fields involved in interoceptive-processing. However, we didn't include these studies because they were generally confounded with exteroceptive signals such as tactile (Komisaruk et al., 2011) or visual stimulation (Parada et al., 2018).

1.2. Experiments on exteroceptive-processing

Based on our definition of exteroceptive-processing, the representation of our body via external sensory signals involves primarily exteroceptive modalities such as visual, auditory and tactile. Therefore, we targeted our article search on four sets of tasks: own face recognition, own body recognition, self-agency, and body ownership. These experiments have been proposed to dissociate one's physical existence

from the external environment, and were discussed within the context of "physical-self" in previous review studies (Gillihan and Farah, 2005; Hu et al., 2016).

Images of our own face and body are the two most distinctive and representative bodily-self signals (Blanke, 2012; Blanke et al., 2015; Ferrè et al., 2014; Hu et al., 2016; Tsakiris, 2008), which are also two of the most extensively studied self-related stimuli (Devue et al., 2007; Kruse et al., 2016; Platek et al., 2006; Platek and Kemp, 2009; Sugiura et al., 2011). In these studies, the participants were required to differentiate between their own photos and photos of another person, which could be a familiar other, e.g. a friend, queen of England, etc., or a total stranger.

While the face and body tasks involve only visual signals, our bodily-self often processes signals from multiple modalities and has to integrate them into a coherent sense of ownership (Seth, 2013). The primarily used task in differentiating self-other ownership of the body is the rubber hand illusion task. In this task, participants are induced with a switch in the sense of ownership from their own hand, which is visually hidden but tactilely stimulated, into a prosthetic hand which is visually seen and synchronously touched (Tsakiris, 2017).

Another important task testing integration of multiple bodily signals of the self is agency. This task investigates the brain's ability in the primary distinction of bodily signals between actively initiating a motion (self-agency) and being the passive recipient of such a motion (other-agency) (Sperduti et al., 2011). In a typical agency task, participants could be required to perform a simple gesture (e.g. flipping of his/her hand), or more commonly, to control an object (e.g. a joystick), meanwhile receiving congruent or incongruent sensory feedbacks from the motion, based on which they were required to judge whether the motion was initiated by themselves or by external causes (Nahab et al., 2011; Powell et al., 2009; Sperduti et al., 2011).

1.3. Experiments on mental-self-processing

For Mental-self-processing, considering the heterogeneity of stimuli involved, we focused our article search on the following tasks: self-other trait adjective / sentence / paragraph judgement, own name recognition, autobiographical memory, object assignment and first/third person perspective judgement.

Trait adjectives are one of the most commonly adopted approaches in investigating social-self attributes. A standard design involves participants judging whether a trait adjective (e.g. honest) is more related to himself/herself, or to another person, which could be a total stranger, a familiar other (e.g. a friend) or a public figure (e.g. queen of England) (Buuren et al., 2010). Another similar, but more flexible form would be to use sentences or paragraphs depicting a certain quality, situation or personality, and participants are required to decide whether the sentences or paragraphs are more relevant to himself/herself or another person (Modinos and Ormel, 2009).

One's own name is a special kind of self-attribute that is also frequently tested in self researches. The name is only an abstract symbol yet it incorporates everything pertaining to the self (Taciowski et al., 2020). Hearing one's own name has been found to invoke special brain responses in EEG signals (Lechinger et al., 2016), an effect prominent in even patients with severe disorders of consciousness (Kempny et al., 2018). In our search, both aural and visually presented own name recognition tasks are included.

Our selves are also shaped by our past. Autobiographical memory tasks usually require participants to differentiate stimuli that are related to their own past events or other people's (Summerfield et al., 2009). Unlike the above-mentioned tasks that rely mainly on semantic processing, autobiographical memories usually involve episodic memories, therefore is often tested via visual stimuli, e.g. photos (Cabeza et al., 2004; Herold et al., 2015).

Experiments have found that our selves can also be extended by our possessions (Belk, 1988; Kim and Johnson, 2014). Using object

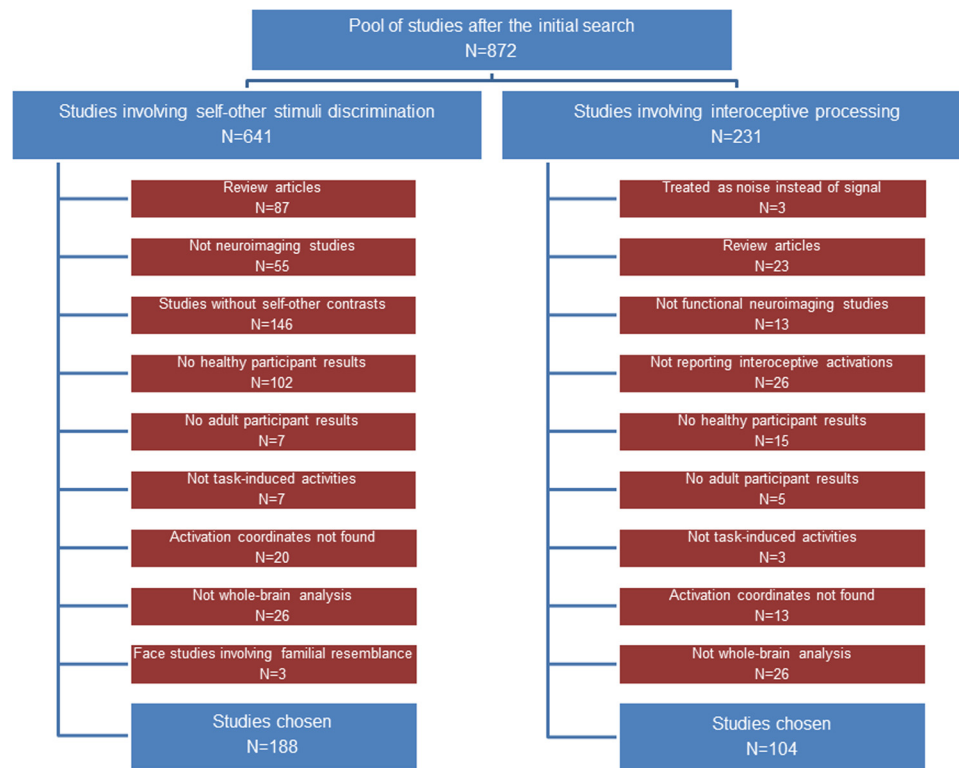


Fig. 2. The PRISMA flow chart of the article selection process. Red color indicates exclusion criterion.

assignment tasks, it is found that random objects – even meaningless ones such as a geometric figure – when assigned to the self in an experimental setting, would automatically acquire higher preference ratings and memory advantage (Kim and Johnson, 2014; Turk et al., 2011), which exhibits differential neural activities than objects assigned to another person (Kim and Johnson, 2014).

During social interactions, we constantly need to imagine “putting ourselves in others’ shoes”. The ability to quickly switch between one’s own view from others’ or vice versa, is crucial in a social life, and is usually tested using the perspective taking tasks. These tasks usually require the participants to differentiate between their own point of view (first person perspective, 1PP) from the point of view of another person (third person perspective, 3PP) under the same scenario, such as a simulated room with various numbers of dots on the walls, or a semantically depicted social event (Ruby and Decety, 2004; Vogeley and Fink, 2003).

1.4. Aims of this study

The general aim of our study is construct a hierarchical neural model of self, which conveys information in three levels to achieve the integration of bodily and external environment information for the self. For these purposes, we used the activation likelihood estimation (ALE) approach to conduct quantitative meta-analyses on different neuroimaging datasets featured by experimental tasks involved in the three processing levels respectively: heartbeat, hunger, thirst, gastric distension, dyspnea/hyperpnea and micturition for Interoceptive-processing; face and body recognition, body ownership and agency for Exteroceptive-processing; and trait adjective / sentence /paragraph judgement, autobiographical memory, name recognition, object assignment and person perspective judgement for Mental-self-processing. Subsequently, the single results from each of the three levels were further compared with overlap analyses and contrast analyses for their differential and common patterns. Based on the results of these analyses, we would eventually try to build a theoretical model for the

neural mechanism of self. Additionally, familiarity has always been an important confounding factor in self-studies, for many argue that the self might be nothing more than extreme familiarity (Gillihan and Farah, 2005; Kircher et al., 2001; Nevi et al., 2016; Qin et al., 2012; Seger et al., 2004; Sugiura et al., 2005). In order to control for this factor, in this study, we also performed a meta-analysis for studies investigating neural activities when processing stimuli related to a personally familiar other (e.g. a close friend, mother), and compared that with the results of the self.

2. Methods and materials

2.1. Literature search and study selection

Articles included in this meta-analysis were collected from a literature search on Pubmed, which was concluded by February, 2020. Search terms include keywords on general or specific aspects of self-processing, such as “self”, “interoception”, “own face”, etc. All search terms were combined with “fMRI” or “PET” to include only functional neuroimaging studies. For specific terms and procedures in the search in each level, see chapters 2.1.1–2.1.4.

To ensure thorough coverage, reference lists of the resulting papers from the above search, and recently published meta-analyses on the self are also searched for missed papers (Araujo et al., 2013; Burrows et al., 2016; Denny et al., 2012; Hu et al., 2016; Kim, 2012; Legrand and Ruby, 2009; Martinelli et al., 2013; Murray et al., 2012; Platek et al., 2008; Qin and Northoff, 2011; Sperduti et al., 2011; van der Meer et al., 2010; van Veluw and Chance, 2014).

After the initial search, a total of 872 articles were found. A screening procedure was then performed using criterion as follows:

- 1 Only whole brain analyses were included. Analyses based on a priori ROIs were excluded.
- 2 Only healthy subjects were included. Patient results were excluded.
- 3 Only studies providing activation coordinates reported in a standard

Table 1
Summary of tasks included in the various analyses.

Classification	Experimental design	Number of articles
Interoceptive-processing	Heartbeat detection	10
	Heartrate variability analysis	17
	Food vs. nonfood / attractive food vs. unattractive food	24
	Hunger vs. satiety state	5
	Thirst vs. satiety state	7
	Micturition	19
	Dyspnea	9
	Hyperprnia	3
	Gastric distension	10
	Exteroceptive-processing	Self/other face discrimination
Self/other body discrimination		6
Self-initiated/other-initiated motion		21
Body ownership & rubber hand illusion		9
Mental-self-processing	Trait adjective judgement	58
	Self-relevance sentence judgement	23
	Person perspective judgement	21
	Name recognition	8
	Object assignment	5
	Autobiographical memory	5
	Self/other judgement in other complex tasks (e.g. mini gambling games or videos)	5
	Familiarity	Face/body-recognition
Name recognition		5
Trait adjective judgement		11
Other tasks (perspective taking, voice recognition, etc.)		5

space (MNI or Talairach) were included.

- 4 Only adult studies were included. Studies on children (aged between 0–17) were excluded.
- 5 Only task induced activation results were included, resting-state functional connectivity results were excluded.

After the screening process, a total of 292 studies on the self and 35 studies on familiarity were included in the final analyses. See Fig. 2 for the detailed article selection procedure, and Table 1 for a summary of the tasks included in the final self-analyses.

2.1.1. Interoceptive-processing

For the Interoceptive-processing level, studies using one of the following tasks were collected: heartbeat, hunger, thirst, dyspnea/hyperprnia, gastric distension and micturition. Search terms used were: “heartbeat” / “heartbeat counting” / “heartrate” / “heartrate variability” / “HRV” / “hunger” / “food” / “thirst” / “in” / “air hunger” / “dyspnea” / “hyperprnia” / “micturition” / “bladder distention” / “interoception”. Two recent meta-analysis studies were referenced for missed papers (Harvie et al., 2019; Schulz, 2016). After screening, eventually a total of 104 studies were selected. For details of all the selected studies at this level, please see Supplementary Table 1.

2.1.2. Exteroceptive-processing

For the exteroceptive-processing level, studies of interest are task-based experiments comparing different brain activation patterns between own-face/body recognition and other-face/body recognition, as well as agency experiments comparing self-initiated and other-initiated motions, and body ownership experiments involving the shift in sense of ownership of the own body parts, e.g. rubber hand illusion. Search terms used were: “self face” / “own face” / “self body” / “own body” / “agency” / “self-agency” / “active passive” / “agent” / “body ownership” / “rubber hand illusion”. After screening, eventually a total of 63 studies were selected. For details of all the selected studies at this level, please see Supplementary Table 2

2.1.3. Mental-self-processing

For the Mental-self-processing level, studies of interest are task-based experiments comparing different brain activation patterns between self-related information and other-related information or a baseline condition. Search terms used were: “self” / “self other” / “self-reference” / “self-referential” / “self name” / “own name” / “self trait” / “trait adjective” / “autobiographical memory” / “perspective taking” / “first perspective” / “1 P P” / “3 P P”. Search results cover a wide range of tasks including: person perspective judgement, own/other name detection, personality trait adjective / sentence / paragraph judgement, autobiographical memory, etc. After screening, eventually a total of 125 studies were selected. For details of all the selected studies at this level, please see Supplementary Table 3.

2.1.4. Familiarity

For the Familiarity condition, same tasks used in 2.1.2 and 2.1.3 were also included. Studies of interest are task-based experiments comparing different brain activation patterns between information of personally familiar people (e.g. a close friend) and personally unfamiliar people (e.g. a stranger), or between information of personally familiar people and self. Search terms used were: “familiarity” / “self familiar” / “self familiarity” / “familiar other”. Search results include tasks of face/body picture recognition, name detection, trait adjective/sentence judgement, etc. In order to control for the confounding effect of familial resemblance on the face, three studies were excluded because they were based on photos of family members who are related to the subject by blood (e.g. father) (Harada et al., 2010; Ma et al., 2014; Wang et al., 2012). After screening, eventually a total of 35 studies were selected in the familiarity condition analysis. For details of the selected studies, please see Supplementary Table 4.

2.2. Overview of the ALE method

We adopted the ALE approach implemented in GingerALE 2.3.6 to perform meta-analyses in this study. The ALE approach, which stands for activation likelihood estimation, is a widely used, automated quantitative approach for voxel-wise neuroimaging meta-analyses (Laird et al., 2005; Turkeltaub et al., 2002). It was first designed by Turkeltaub in 2002, and has been adopted by Brainmap® since 2003. This algorithm calculates the probabilities for foci (i.e. coordinates for maximum activations) reported in the studies to be “true” convergence across different laboratories, reflecting meaningful mental operations rather than random clustering in the brain (Eickhoff et al., 2009; Turkeltaub et al., 2002). The key idea is that, due to spatial uncertainties of reported activation coordinates inherent in functional neuroimaging studies, foci reported from individual studies should not be considered as points, but as distribution probabilities representing the likelihood of activation for voxels surrounding each focus (Turkeltaub et al., 2002).

In order to calculate the distribution probability for each focus, a 3D Gaussian function is applied so that the probability of a given focus lying within a voxel is:

$$p = \frac{e^{-d^2/2\sigma^2}}{(2\pi)^{1.5}\sigma^3}$$

where d is the Euclidean distance between the center of the voxel and the focus, and σ is the standard deviation of the distribution. A “Modelled Activation” map (i.e. MA map) is thus constructed for each study representing all foci reported in that study, so that a meta-analysis on 10 studies, for instance, will generate 10 MA maps. Combining all the MA maps for all the studies within an analysis, an ALE map was then created by merging all the MA scores into ALE scores in a single ALE map, representing the probabilities of observing activation from at least one study in each voxel. To draw inference from the analysis, a null distribution map is constructed by randomly drawing the same

Table 2
Supra-threshold clusters in three levels of self-processing analysis.

#	Hemi	Label	Brodman Area	Volume (mm ³)	Peak Z Value	x	y	z
Interoceptive-processing								
1	R	Insula	BA13	9656	5.79	34	14	12
2	L	Dorsal anterior cingulate cortex	BA24	5768	5.63	0	4	48
3	R	thalamus		5408	6.12	12	-14	4
4	R	Parahippocampal Gyrus		1416	4.33	30	-4	-24
5	L	Parahippocampal Gyrus		1216	5.71	-20	-4	-20
6	L	Insula	BA13	1000	4.63	-40	-2	2
7	L	Insula	BA13	976	4.30	-36	24	4
8	R	Inferior parietal lobule	BA40	728	4.37	56	-26	26
9	R	Superior frontal gyrus	BA8	376	4.29	4	24	48
10	L	Superior temporal gyrus	BA22	352	3.86	-56	6	6
11	L	Postcentral gyrus	BA2	232	3.91	-48	-16	32
Exteroceptive-processing								
1	R	Fusiform gyrus	BA37	2544	6.25	48	-58	-12
2	R	Inferior frontal gyrus	BA46	2520	6.72	48	40	8
3	R	Premotor cortex	BA9	2336	7.41	50	8	26
4	R	Insula	BA13	2216	5.48	40	8	0
5	L	Fusiform gyrus	BA19	1720	6.46	-44	-68	-6
6	R	Superior parietal lobule	BA7	1712	5.39	26	-72	44
7	R	Postcentral gyrus	BA3	1328	5.95	58	-22	38
8	R	Inferior parietal lobule	BA40	1320	4.67	36	-50	56
9	L	Insula	BA13	856	4.28	-36	18	-4
10	R	Inferior occipital gyrus	BA19	456	4.59	38	-80	-2
11	L	Inferior parietal lobule	BA40	360	3.88	-46	-34	40
12	L	Superior parietal lobule	BA7	248	3.79	-22	-64	50
13	R	Cingulate gyrus	BA32	240	4.01	4	8	38
14	L	Medial prefrontal cortex	BA10	216	4.18	-6	60	22
Mental-self-processing								
1	L	Anterior Cingulate cortex / medial prefrontal cortex	BA32/BA10	17,616	8.31	-6	48	0
2	L	Posterior cingulate cortex	BA31	4704	6.05	-4	-54	28
3	L	Insula	BA13	2584	4.89	-36	22	-2
4	L	Middle temporal gyrus	BA39	2240	5.06	-48	-66	28
5	L	Thalamus		984	4.35	-8	2	8
6	L	Superior frontal gyrus	BA8	792	4.11	-20	36	46
7	L	Cingulate gyrus	BA24	768	4.92	0	-18	40
8	L	Inferior temporal gyrus	BA21	480	4.22	-62	-6	-18
9	R	Middle temporal gyrus	BA39	416	4.49	54	-60	24
10	R	Insula	BA13	288	3.99	52	10	-6
11	L	Superior frontal gyrus	BA10	280	3.70	-24	50	22
12	R	Premotor cortex	BA6	256	3.88	46	6	24

Note: Thresholds were set at $p < 0.05$, FDR corrected, with a minimum cluster size of 200 mm³. **Abbreviations:** Hemi—hemisphere; L—left; R—right; BA—Brodman area.

number of foci as in the ALE map and applying a Monte Carlo permutation approach, from which all voxel values were collected into a histogram. The values in the histogram were then used to form the null hypothesis, with which a significance threshold can be established, so that values from the “real” ALE map exceeding that threshold would represent with confidence that a true convergence among studies, rather than random clustering, is found (Turkeltaub et al., 2002).

2.3. Statistical analyses

2.3.1. Single analyses

Coordinates were manually collected from the papers and coded into 4 conditions according to criteria described in chapters 2.1.1–2.1.4. If multiple contrasts were reported in one experiment, only the one with the highest self-specificity would be used. For instance, if an experiment reported results from both self vs baseline and self vs other contrasts, only the self vs other contrast would be chosen. Coordinates originally reported in TLRC space were transformed into MNI space using the icbm2tal tool (Lancaster et al., 2007) implemented in GingerALE 2.3.6. The FDR approach was applied to correct for multiple comparisons. In our single analyses for each condition, clusters were thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm³ (Burrows et al., 2016; Martinelli et al., 2013; van Veluw and Chance, 2014).

Additionally, to differentiate self from familiarity, we also

constructed a (general) Self-analysis as a matched condition for the later contrast analysis between self and familiarity, combining all studies from 2.1.2 and 2.1.3. For details as well as results of this Self-analysis, please refer to Supplementary Fig. 1 and Supplementary Table 5.

2.3.2. Overlap analysis of the different levels of self-processing

To identify the primary regions for self-processing, after obtaining single results from analyses of each level of self-processing, we further performed an overlap analysis, superimposing the three results to find the regional overlap. Because exteroceptive-processing is hypothesized to be the link between bodily and external environment information, there should be regions seen at this level of processing that overlap with the other two levels, which are likely to be where non-bodily environment information is integrated with interoceptive information. Moreover, the overlap between self and familiarity was also analyzed to see whether and how the two are similar. These analyses were conducted using AFNI. Using the 3dcalc program, we converted all non-zero ALE values in each single image into 1, and multiplied the values in corresponding voxels from different images, so that only clusters with non-zero values in every image analyzed could survive.

2.3.3. Contrast analyses

In this study, we're interested to know not only what brain regions are involved in each level of self-processing, but also how they change

Table 3
Supra-threshold clusters in the contrast analyses between Interoceptive-processing, Exteroceptive-processing and Mental-self-processing.

#	Hemi	Label	Brodman Area	Volume (mm ³)	Peak Z Value	x	y	z
Interoceptive-processing > Exteroceptive-processing								
1	R	Thalamus		3656	3.67	17	-18	3
2	L	Cingulate gyrus	BA24	3568	3.78	-1	-2	45
3	R	Insula	BA13	3024	3.49	30	11	14
4	R	Parahippocampal gyrus		1400	4.11	21	-8	-23
5	L	Parahippocampal gyrus		1216	4.11	-20	-6	-19
6	R	Inferior parietal lobule	BA40	384	3.03	56	-32	26
7	R	Inferior frontal gyrus	BA44	344	1.91	58	8	4
8	L	Insula	BA13	304	2.18	-42	-4	6
9	R	Inferior frontal gyrus	BA4	272	2.69	62	-4	18
10	L	Postcentral gyrus	BA2	224	2.09	-46	-20	32
Exteroceptive-processing > Interoceptive-processing								
1	R	Fusiform gyrus	BA37	2168	4.11	50	-54	-13
2	R	Premotor cortex	BA9	2064	3.94	52	9	27
3	R	Middle frontal gyrus	BA46	2048	3.94	48	38	8
4	R	Superior parietal lobule	BA7	1672	3.67	28	-64	46
5	L	Inferior temporal gyrus	BA37	1664	4.11	-46	-70	-1
6	R	Superior parietal lobule	BA7	1184	3.11	34	-54	66
7	R	Postcentral gyrus	BA2	1056	3.26	58	-22	44
Exteroceptive-processing > Mental-self-processing								
1	R	Premotor cortex	BA9	4992	3.89	51	9	28
2	R	Middle frontal gyrus	BA46	2744	3.89	45	40	10
3	R	Fusiform gyrus	BA37	2408	3.89	49	-58	-7
4	R	Precuneus	BA7	1840	3.89	26	-61	48
5	L	Fusiform gyrus	BA37	1680	3.89	-46	-64	-14
6	R	Temporal parietal junction		1320	3.89	52	-26	38
7	R	Superior parietal lobule	BA7	1320	3.89	39	-48	52
8	R	Fusiform gyrus	BA37	304	2.64	34	-82	-6
9	L	Temporal parietal junction	BA40	288	2.66	-48	-38	36
10	L	Precuneus	BA7	248	3.72	-24	-62	47
11	R	Middle cingulate gyrus		232	2.79	4	8	42
Mental-self-processing > Exteroceptive-processing								
1	L	Pregenuan anterior cingulate / medial prefrontal cortex	BA32 / BA10	11,192	3.89	-7	53	-1
2	L	Posterior cingulate cortex	BA7/BA31	3280	3.89	-7	-50	36
3	L	Temporal parietal junction	BA39	744	2.18	-44	-54	32
4	L	Caudate		328	2.31	-8	6	16
5	L	Superior frontal gyrus	BA8	248	2.11	-20	34	46
6	L	Superior frontal gyrus	BA10	224	2.60	-22	52	22
Interoceptive-processing > Mental-self-processing								
1	R	Insula	BA13	8744	2.26	38	5	9
2	L	Cingulate Gyrus	BA24	5744	3.94	-1	5	42
3	R	Thalamus		4936	3.94	8	-13	4
4	R	Parahippocampal Gyrus		1416	4.11	26	-4	-21
5	L	Parahippocampal Gyrus		1168	4.11	-23	-3	-18
6	L	Insula	BA13	784	2.78	-40	4	2
7	R	Inferior Parietal Lobule	BA40	736	4.11	53	-26	29
8	R	Clastrum		432	3.45	39	8	-13
9	L	Precentral Gyrus	BA6	200	2.20	-48	-12	34
Mental-self-processing > Interoceptive-processing								
1	L	Medial Frontal Gyrus	BA9	12,288	4.11	-4	57	-5
2	L	Posterior Cingulate cortex	BA31	2144	3.67	-1	-57	29
3	L	Middle Temporal Gyrus	BA39	1952	3.49	-44	-62	28
4	L	Superior Frontal Gyrus	BA8	640	3.10	-16	40	44
5	R	Cingulate Gyrus	BA24	576	2.87	4	-18	42
6	L	Inferior Temporal Gyrus	BA21	456	4.11	-60	-10	-19
7	L	Caudate		456	2.66	-8	4	14
8	R	Superior Temporal Gyrus	BA39	416	3.33	56	-58	22

Note: Thresholds were set at $p < 0.05$, with a minimum cluster size of 200 mm³. **Abbreviations:** Hemi—hemisphere; L—left; R—right; BA—Brodman area.

between levels. For that purpose, we combined different approaches, i.e. single analysis, overlap analysis and contrast analysis. In a hierarchical system, ALE single analyses could help find brain regions underlying a certain condition, and overlap analyses reveal the fundamental region across different conditions. Contrast analyses on the other hand, reveal the differential brain regions between two conditions, thus indicate how information propagates from one level to another. Therefore, by combining the three approaches, we could find the neural pathways through which information propagates from the internal body to the external environment in the three levels of self-processing.

For these purposes, we performed contrast analyses on paired single

results from the previous steps, using the subtraction method in GingerALE 2.3.6. The following contrasts were made: Interoceptive-processing vs. Exteroceptive-processing, Exteroceptive-processing vs. Mental-self-processing, Interoceptive-processing vs. Mental-self-processing, and self vs. familiarity. These contrasts compare the difference of ALE scores between each pair of images using permutation tests, so that a significant result indicates a cluster of voxels having significantly higher probability of observing activations reported from studies in one condition with respect to the other. Therefore, a single analysis reveals brain regions that are highly relevant in a certain condition, and a contrast analysis reveals brain regions that are more specific to this condition with respect to another.

Table 4
Supra-threshold clusters in the Familiarity condition, as well as the contrast analysis between self and familiarity.

#	Hemi	Label	Brodmann Area	Volume (mm ³)	Peak Z Value	x	y	z
Familiarity								
1	R	Posterior cingulate cortex	BA31	3320	5.55	6	−54	28
2	R	Ventromedial prefrontal cortex	BA10	2208	5.03	6	56	4
3	L	Anteromedial prefrontal cortex	BA9	872	5.42	−6	56	20
4	L	Dorsomedial prefrontal cortex	BA8	432	4.73	−6	52	40
5	L	Inferior frontal gyrus	BA47	384	4.63	−48	26	−8
6	L	Temporal parietal junction	BA39	280	4.34	−48	−66	34
7	R	Middle temporal gyrus	BA21	208	3.99	64	−8	−16
Self > Familiarity								
1	R	Precentral gyrus	BA6	2088	2.56	46	2	30
2	R	Middle frontal gyrus	BA46	856	2.23	44	36	10
3	R	Insula	BA13	336	1.93	43	3	−1
4	R	Fusiform gyrus	BA19	328	2.04	55	−61	−3
5	R	Precentral gyrus	BA6	320	2.47	50	4	38
6	R	Precuneus	BA7	224	2.30	24	−74	46
7	L	Middle occipital gyrus	BA37	200	1.96	−48	−72	−2
Familiarity > Self								
1	R	Posterior cingulate cortex	BA31	3976	3.89	6	−57	29
2	R	Anteromedial prefrontal cortex	BA9	2904	3.04	4	58	8
3	L	Inferior frontal gyrus	BA47	648	2.73	−48	24	−4
4	L	Dorsomedial prefrontal cortex	BA8	544	3.89	−5	51	38
5	R	Dorsomedial prefrontal cortex	BA32	384	2.79	2	28	32
6	L	Temporal parietal junction	BA39	352	2.93	−46	−64	38
7	R	Inferior temporal gyrus	BA21	208	3.54	65	−8	−17

Note: In the Familiarity condition, thresholds were set at $p < 0.05$, FDR corrected, with a minimum cluster size of 200 mm³. In Self > Familiarity and Familiarity > Self, thresholds were set at $p < 0.05$, with a minimum cluster size of 200 mm³. **Abbreviations:** Hemi—hemisphere; L—left; R—right; BA—Brodmann area.

The contrast analyses were performed in four steps:

- 1) A pair of single images were constructed using the methods described in 2.3.1, thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm³. A mask was then created to include all significant regions in the single datasets, to limit the calculations within the mask.
- 2) Simple subtractions were done between the paired images, subtracting the ALE scores (in both directions) between each pair of voxels in the two images, creating an A–B and a B–A difference map;
- 3) Permutation tests were performed to create simulated difference scores for statistical inferences to be made. Firstly, a pooled dataset was constructed by combining the A and B datasets and analyzed in the same way as in 2.3.1. The pooled dataset was then divided randomly into two simulated datasets with the same sample sizes as in the original data. The simulated data were subtracted with each other, creating simulated difference maps which were compared with the true difference maps. This process was iterated for 25,000 times, from which a P value was calculated for each voxel representing where the true score sits on the distribution of values in that voxel (Hu et al., 2016; Laird et al., 2005).
- 4) Inferences were made based on the significance of the true difference scores according to the P value map calculated in step 3. A threshold of $p < 0.05$ was adopted, with an additional cluster volume threshold of 200mm³. The P value map was converted into Z value maps for the final output (Fisher, 1915; Laird et al., 2005).

Visualization of images was presented with Mango (Downloaded from <http://brainmap.org>). Images were overlaid on a standard brain in an MNI template downloaded from the same website. Anatomical labelling was done with reference to the Nearest Grey Matter MNI labels provided by Mango.

3. Results

3.1. Single analyses of the three levels of self-processing

For visual demonstration of the single results, please see Fig. 3. Cluster details are listed in Table 2.

3.1.1. Interoceptive-processing

A total of 1340 foci generated from 104 studies were analyzed. The result shows 11 contiguous clusters. As shown in Fig. 3-A, three of the clusters were located at the insula, including two at the left anterior insula and one at the right insula. Other regions mainly include the dorsal anterior cingulate cortex (dACC), thalamus and bilateral parahippocampus gyrus.

3.1.2. Exteroceptive-processing

A total of 508 foci generated from 63 studies were analyzed. The result revealed 14 contiguous clusters. As shown in Fig. 3-B, primary findings include the left anterior insula, right middle insula, anteromedial prefrontal cortex (AMPFC), premotor cortex (PMC) and bilateral temporal parietal junction (TPJ). Other clusters include typical face-recognition regions such as right fusiform gyrus, and sensorimotor areas such as the postcentral gyrus.

We've also conducted two separate analyses to verify our assumption that these experiments share great some commonalities and could be analyzed together. One is on Self-recognition, which included the self-face and self-body recognition tasks, and the other is on Agency-ownership, which included the self-agency and body-ownership tasks. The results showed that the two conditions have overlaps in multiple regions, please see Supplementary Fig. 2.

3.1.3. Mental-self-processing

A total of 961 foci generated from 126 studies were analyzed, and 12 clusters were identified. As shown in Fig. 3-C, the results primarily include bilateral insula, as well as pregenual anterior cingulate cortex (pACC) / AMPFC, posterior cingulate cortex (PCC), PMC and bilateral TPJ.

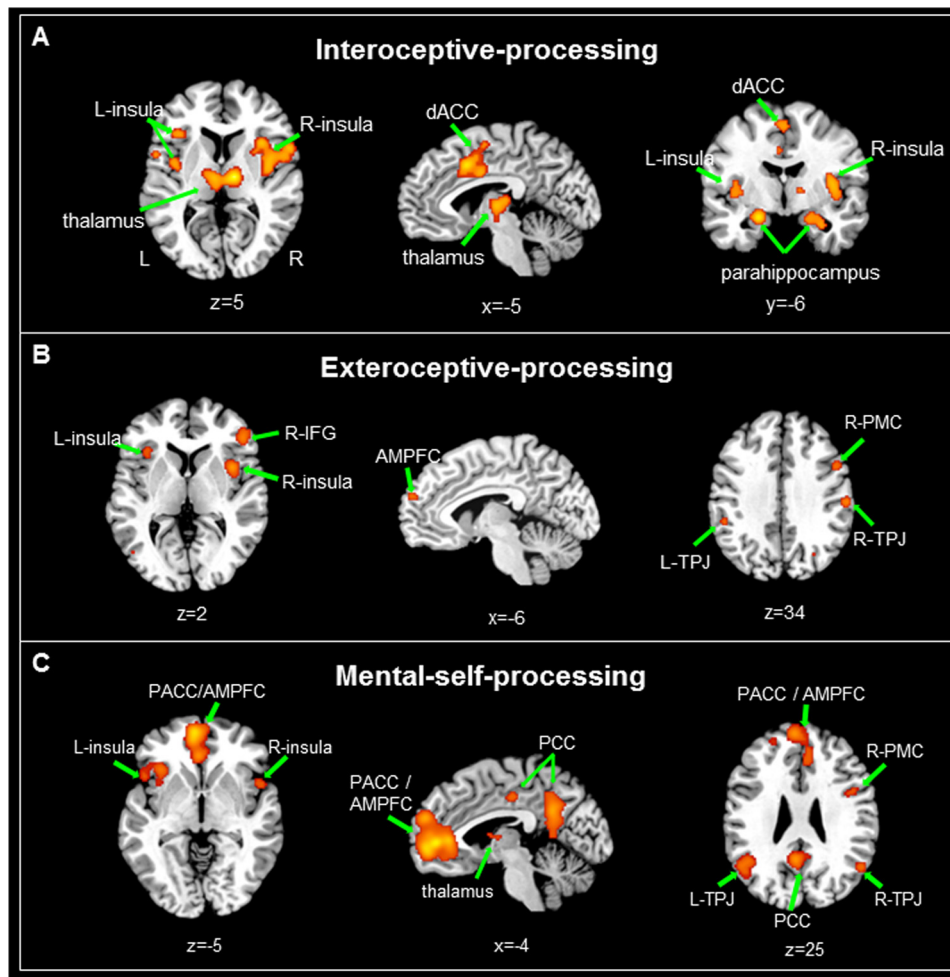


Fig. 3. Single analysis results of the three levels of self-processing. **A:** Interoceptive-processing; **B:** Exteroceptive-processing; **C:** Mental-self-processing. Clusters are thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm^3 . **Abbreviations** – L: left; R: right; dACC: dorsal anterior cingulate cortex; IFG: inferior frontal gyrus; TPJ: temporal parietal junction; AMPFC: anteromedial prefrontal cortex; pACC: pregenual anterior cingulate cortex; PCC: posterior cingulate cortex; PMC: premotor cortex.

3.2. Overlapping of the three levels of self-processing

For illustrative purposes, we superimposed the three single analysis results onto a single brain to show the overlapped regions of the different levels of information processed by the self. As seen in Fig. 4, the overlap is at the left anterior insula, and partially at the right middle insula. Apart from that, overlaps at AMPFC and PMC, as well as a common activation of two clusters in the left TPJ areas, are also seen in Exteroceptive-processing and Mental-self-processing. For cluster information, please see Supplementary Table 5.

3.3. Contrast analyses among the three levels of self-processing

For visual demonstration of the contrast analysis results, please see Fig. 5. Cluster information can be found in Table 3.

3.3.1. Interoceptive-processing vs. Exteroceptive-processing

Ten clusters survived the analysis of Interoceptive-processing > Exteroceptive-processing. As shown in the left panel of Fig. 5-A, results include the right anterior insula and posterior insula, thalamus and bilateral parahippocampus. Seven clusters survived the analysis of Exteroceptive-processing > Interoceptive-processing. As shown in the right panel of Fig. 5-A, results primarily include the right mid-insula, right TPJ, right PMC and right inferior frontal gyrus (IFG).

3.3.2. Exteroceptive-processing vs. Mental-self-processing

11 clusters survived the analysis of Exteroceptive-processing > Mental-self-processing. As shown in the left panel of Fig. 5-B, the results primarily include the right insula, right PMC, as well as bilateral TPJ. Six clusters survived the analysis of Mental-self-processing > Exteroceptive-processing. As shown in the right panel of Fig. 5-B, the results primarily include pACC/AMPFC, PCC and left TPJ.

3.3.3. Interoceptive-processing vs. Mental-self-processing

Nine clusters survived the analysis of Interoceptive-processing > Mental-self-processing. As shown in the left panel of Fig. 5-C, results mainly include bilateral insula, dorsal anterior cingulate, and thalamus. Eight clusters survived the analysis of Mental-self-processing > Interoceptive-processing. As shown in the right panel of Fig. 5-C, results primarily include AMPFC, posterior cingulate cortex (PCC) and bilateral TPJ.

3.4. Familiarity

Results from single analysis of Familiarity, contrast analysis of Self vs. Familiarity and overlapping analysis of Self and Familiarity can be seen in Fig. 6 and Table 4. For details of the Self condition used in these analyses, please see Supplementary Fig. 1 and Supplementary Table 6.

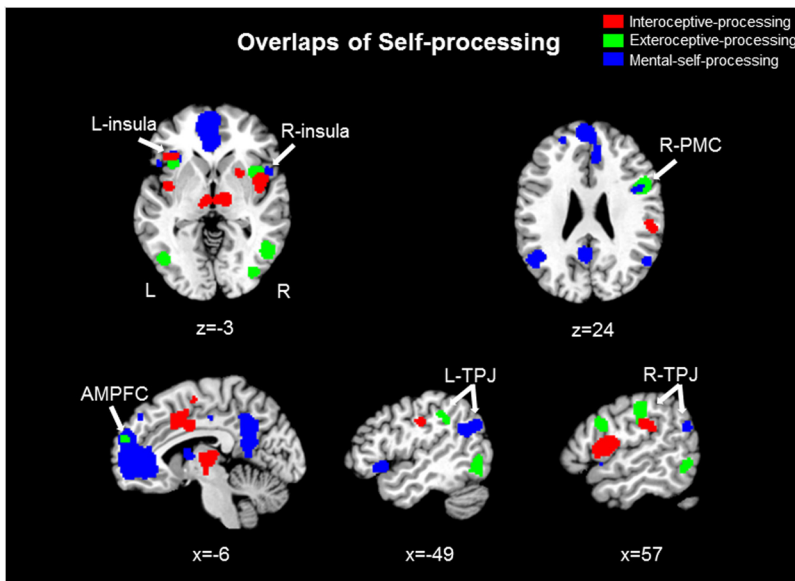


Fig. 4. Overlapping of the three levels of self-processing. Clusters are thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm^3 . Please refer to the legends for colors representing different levels of analysis. **Abbreviations** –AMPFC: anteromedial prefrontal cortex; TPJ: temporal parietal junction; PMC: premotor cortex; L: left; R: right.

3.4.1. Single analysis of Familiarity

A total of 350 foci were analyzed and revealed seven clusters. As shown in Fig. 6-A, the results mainly include the AMPFC, PCC and left TPJ. No clusters in the insula cortex were found.

3.4.2. Self vs. Familiarity

Seven clusters survived the analysis of Self > Familiarity. As shown in the left panel of Fig. 6-B, the results primarily include the right insula, IFG and right PMC. Seven clusters survived the analysis of Familiarity > Self. As shown in the right panel of Fig. 6-B, the results primarily include midline regions such as the AMPFC, dorsomedial

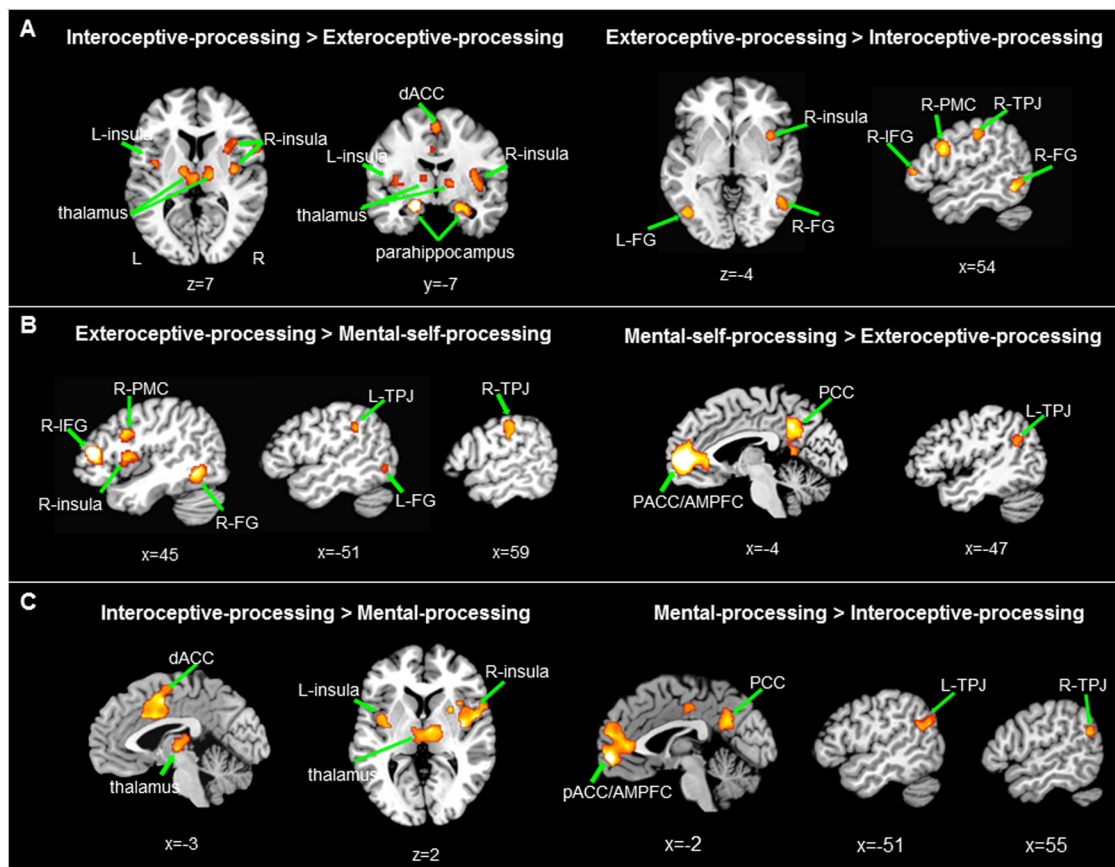


Fig. 5. Contrast analysis results of the three levels of self-processing. **A:** Interoceptive-processing vs. Exteroceptive-processing. **B:** Exteroceptive-processing vs. Mental-self-processing. **C:** Interoceptive-processing vs. Mental-processing. Clusters are thresholded at $p < 0.05$, with a minimum cluster size of 200mm^3 . **Abbreviations** – IFG: inferior frontal gyrus; TPJ: temporal parietal junction; dACC: dorsal anterior cingulate cortex; PMC: premotor cortex; PCC: posterior cingulate cortex; pACC: pregenual anterior cingulate cortex; AMPFC: anteromedial prefrontal cortex; FG: fusiform gyrus; L: left; R: right.

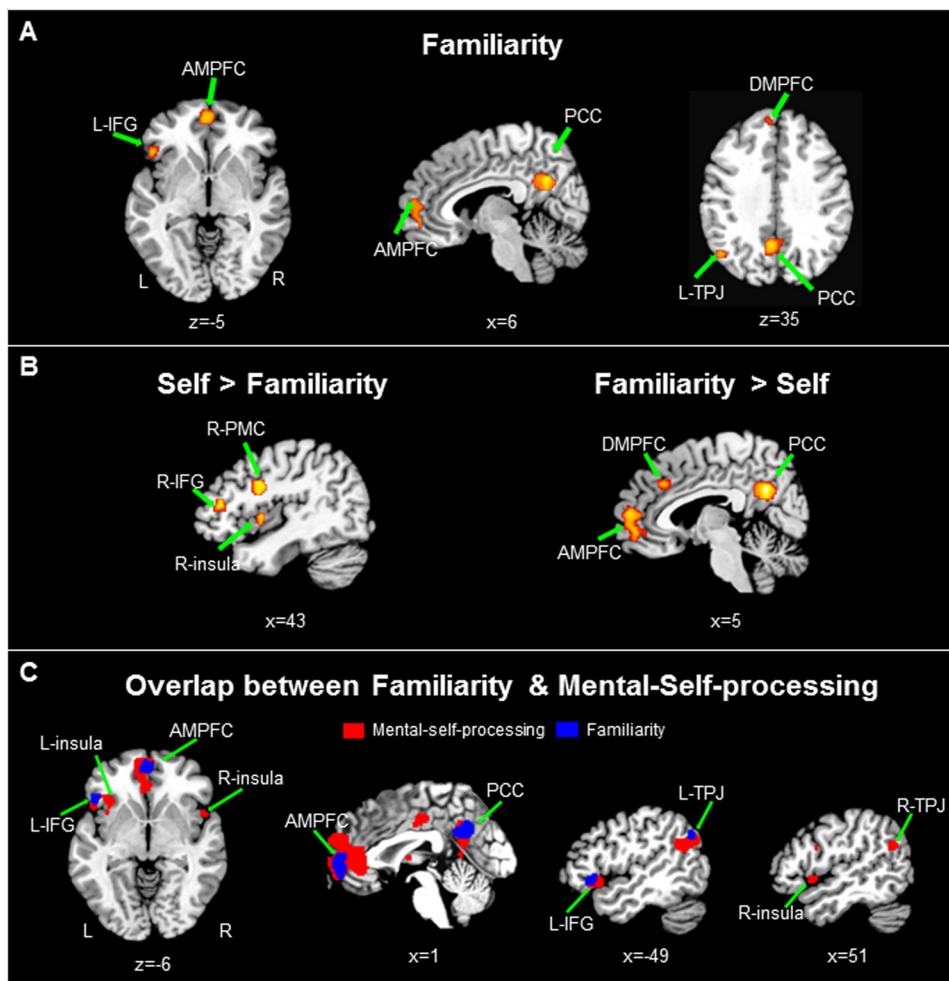


Fig. 6. Familiarity analyses results. **A:** Single analysis result for Familiarity. Clusters are thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm^3 . **B:** Contrast analysis of Self vs. Familiarity. Clusters are thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm^3 . **C:** Overlapping of Self and Familiarity. Clusters are thresholded at $p < 0.05$, FDR corrected, with a minimum cluster size of 200mm^3 . **Abbreviations** – IFG: inferior frontal gyrus; PCC: posterior cingulate cortex; pACC: pregenual anterior cingulate cortex; DMPFC: dorsomedial prefrontal cortex; AMPFC: anteromedial prefrontal cortex; TPJ: temporal parietal junction; PMC: premotor cortex; L: left; R: right.

prefrontal cortex (DMPFC) and PCC.

3.4.3. Overlapping between Mental-self-processing and Familiarity

Overlap analysis is also performed by superimposing the Mental-self-processing and Familiarity results onto a single brain to see the similarity and difference between the two conditions. As seen in Fig. 6-C, highly overlapped regions can be seen in the midline regions at the AMPFC and PCC, as well as left IFG and TPJ. However, the insula activities are only seen in Mental-self-processing, but not in Familiarity.

4. Discussion

In the current manuscript, a three-level processing model of the self was proposed to illustrate how the brain integrates bodily information and external environment information in self-processing. Within this model, Interoceptive-processing represents the internal bodily information which is the prerequisite for the self; Exteroceptive-processing integrates internal and external sensory signals and links the body with external environment information; and Mental-self-processing introduces self-relatedness to external non-bodily stimuli such as personal belongings and social attributes. Overall, information of bodily and external environment is integrated for the self via propagation from Interoceptive-processing to Mental-self-processing.

To explore the neural substrates for each level of self-processing, we here conducted an ALE meta-analysis of neuroimaging studies. A primary finding is the overlap of the bilateral insula across all three levels, while familiarity did not show any consistent activation in the insula. Moreover, the Exteroceptive-processing and Mental-self-processing analyses showed regional overlap in AMPFC, PMC, as well as a common

involvement of the TPJ areas, and also showed more consistent activations in these regions than Interoceptive-processing. Finally, compared with both Interoceptive-processing and Exteroceptive-processing, Mental-self-processing showed more consistent findings in pACC/MPFC and PCC, which are the crucial regions of the default-mode network. These results indicated that from Interoceptive-processing, Exteroceptive-processing to Mental-self-processing, the common regions involved in self-processing extended from the insula, to AMPFC/PMC/TPJ and then to pACC and PCC. These results showed a neural pattern for self-processing that, the more external the information is processed, the more extended the brain regions are involved (see Fig. 7).

4.1. The overlap in the insula across all three levels of self-processing

Our finding of the consistent involvement of the insula in self-processing suggests that its major function, i.e. internal sensory integration / interoceptive-processing, could be the core for the self, which is in line with several researcher's theories (Craig, 2009; Damasio, 2003a; Seth, 2013; Tsakiris, 2017). Based on these theories, interoceptive-processing could provide the foundation for higher-order self-processing.

The insula is primarily found to be an important hub in mediating interoceptive signals (Enzi et al., 2009; Craig, 2009; Wiebking et al., 2015a, 2015b), and different sub-divisions of the insula were found to be involved in different functions. The insula has been proposed to have a posterior-to-anterior progression, representing sensory signals from primary (objective) to higher (subjective) levels (Craig, 2009). Specifically, in this progression, primary (objective) sensory signals

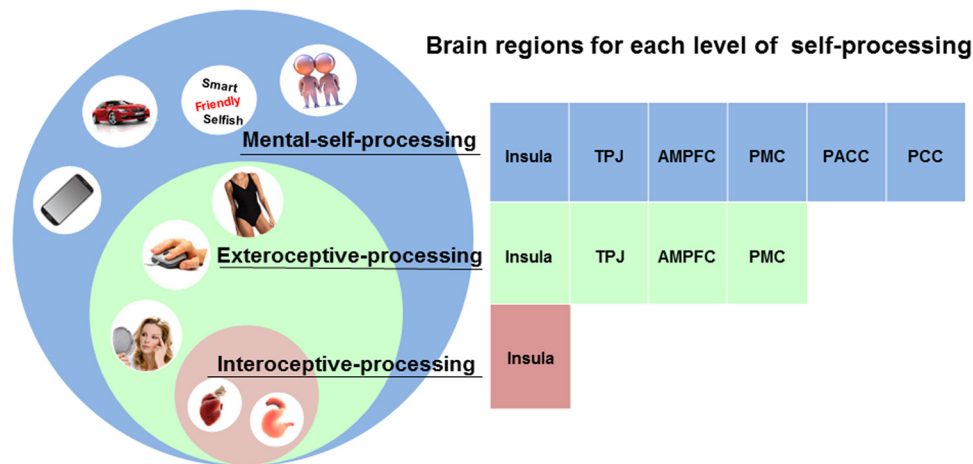


Fig. 7. Brain regions for each level of self-processing. **Abbreviations** – TPJ: temporal parietal junction; AMPFC: anteromedial prefrontal cortex; PMC: premotor cortex; PACC: pregenual anterior cingulate cortex; PCC: posterior cingulate cortex.

ascending from gastrointestinal (as reflected in our hunger and thirst tasks), cardiovascular (as reflected in our heartbeat detection tasks), respiratory and general somatosensory pathways (as reflected in the agency tasks), etc., are converged in the posterior and mid insula (Namkung et al., 2017; Stephani et al., 2011; Uddin et al., 2017), which are then integrated with emotional, cognitive and motivational information from other cortical and subcortical areas such as the ACC and amygdala, and re-represented as subjective feeling states in the anterior insula (Craig, 2009; Critchley et al., 2004).

The introduction of subjective feeling states brought upon by the anterior insula could constitute a neural core for the awareness of a bodily-self (Craig, 2009; Damasio, 2003a; Namkung et al., 2017; Seth, 2013; Enzi et al., 2009). Combined with our finding of its consistent activation in all three levels of self-processing, it is possible that even in the external environment where seemingly minimum bodily-signals are involved, insula activities, i.e. internal sensory information, is still necessary for the external environment signals to be explicitly recognized as self-related signals.

There have been implications of a right lateralized role for the insula in self-processing (Craig, 2009). For instance, one study found that the right anterior insula was specifically involved in own face recognition with respect to familiar face or a scrambled image (Deveue et al., 2007). However, in our current analyses, no obvious laterality can be seen for the insula activations. One possible explanation is that the left and right insula might be specialized in different functions, which cannot be differentiated in self-studies. For instance, the left insula has been proposed to be more involved in positive and negative emotions (Leibenluft et al., 2004), whereas the right insula is more involved in bodily sensations (Craig, 2009), which has been found to be involved in both interoception and exteroception (Simmons et al., 2012). It is possible that the left anterior insula could be responsible for assessing the affective element of self-related information, and the right anterior insula could be responsible for co-representing interoceptive-exteroceptive information. However, more studies should be performed to verify these assumptions.

4.2. AMPFC and TPJ in Exteroceptive-processing and Mental-self-processing

In the Exteroceptive-processing tasks, external environment signals, which have a direct relationship with the body, have to be integrated with interoceptive signals to build the link between the bodily and external environment information (Park and Blanke, 2019), and through this link, non-bodily information from the external environment could be integrated with the interoceptive signals, for the non-

bodily information to become explicitly self-relevant. The regional overlap in AMPFC, TPJ and PMC found between Exteroceptive- and Mental-self-processing, as well as their higher involvement in Exteroceptive-processing than Interoceptive-processing, suggests that these regions are likely to be neural substrates for the realization of such integration or linkage.

4.2.1. AMPFC and the self

Consistent with our findings, AMPFC activities have been found in a wide range of self-processing tasks. In one previous review, the authors used different approaches to compare the three subdivisions of MPFC – the VMPFC (ventromedial prefrontal cortex), AMPFC and DMPFC, in three different functional domains – social cognition, self-processes and valuation, and found that the AMPFC is especially involved in self-processing (Lieberman et al., 2019). Furthermore, lesion studies focusing on AMPFC damage has found that patients exhibit impaired self-referential processing (Kurczek et al., 2015; Philippi et al., 2012), conceptual self-knowledge (Marquine et al., 2016), and self-conscious emotion (Beer et al., 2003). Moreover, another meta-analysis used logistic regression analysis in comparing differential functioning in AMPFC and DMPFC, found that while both AMPFC and DMPFC are involved in both self and other judgements, a stronger involvement in self judgements is seen in AMPFC (Denny et al., 2012).

Although with such strong implications for self-processing in AMPFC, exactly what function does the AMPFC performs within the self is still unclear. One meta-analysis proposed that MPFC serves a function of basic and higher-order differentiation between self and other (van Veluw and Chance, 2014). Based on our results, we further propose that it's the subdivision of MPFC, the AMPFC, which specializes in this function. Furthermore, given its presence in both bodily and non-bodily signals, we suspect that AMPFC's role in self-other differentiation is basic, rather than higher-order. Analogous to the immune system, self-other differentiation is a prerequisite in maintaining the integrity of a living organism, and should therefore be ubiquitous in dealing with the external world (Gonzalez et al., 2011).

Tractography studies have found fibers connecting the AMPFC/pACC with the anterior insula (Ghaziri et al., 2017). Combined with our findings, it is likely that the AMPFC and insula could have collaborated into forming a preliminary distinction between self-related and non-self signals, so as to allow for any further self-external interaction. This explains the absence of AMPFC activity in Interoceptive-processing, since no external information is involved in dealing with interoceptive signals, and therefore self-other differentiation is not necessary. Based on this hypothesis, the collaboration between the AMPFC and insula could be a first step for our internal world to reach out, and a requisite

for an external signal to be perceived as a self-signal.

4.2.2. TPJ and the self

The TPJ areas are also commonly involved in both the Exteroceptive- and Mental-self-processing. The TPJ area is a sparsely interconnected network, functionally defined from a wide variety of higher order functions including sensory, cognitive, emotional, social and motor levels (Eddy, 2016; Igelström and Graziano, 2017). Nodes within this region include IPL, posterior superior / middle temporal gyrus (pSTG/MTG), supramarginal gyrus (SMG) and angular gyrus (AG) (Igelström and Graziano, 2017).

Consistent with our findings of the Exteroceptive-processing analysis, the TPJ activation is frequently found in tasks involving multi-sensory and sensorimotor integration, which is essential for bodily self-consciousness (Lenggenhager et al., 2006; Eddy, 2016). For instance, lesion studies have found that damage to the right TPJ could induce an “out-of-body” experience (Lenggenhager et al., 2006). TMS study also found that disruption to the right TPJ could result in a blurred self-other boundary of one’s own body parts (Tsakiris et al., 2008).

Structural studies have found that the TPJ is connected to the anterior insula via the middle longitudinal fasciculus and extreme capsule (Saur et al., 2008). Combined with our current results, it is strongly implicated that the TPJ is responsible for integrating multi-sensory and sensorimotor signals through collaboration with the insula, and thus constituting a coherent sense of bodily-self (Park and Blanke, 2019). Indeed studies have found that disruption to the connectivity between the TPJ and insula could induce “personal neglect”, in which the individual loses awareness of the contralesional half of their body (Committeri et al., 2007; Friedrich et al., 1998; Halligan et al., 2003).

However, consistent with our results that TPJ was involved in integrating the bodily and external environment information, TPJ is also frequently implicated in the social or mental aspects of the self (Eddy, 2016). For instance, TPJ activities are frequently found in theory of mind (TOM) tasks (Chan and Lavalée, 2015; Schurz et al., 2017) as well as imitation and perspective taking (Schurz et al., 2013). A commonality of these tasks is that they all require an on-line control, i.e. simultaneous representation, of both self and other (Santiesteban et al., 2015, 2012; Schurz et al., 2013). For instance, using transcranial direct current stimulation (tDCS), one research found that excitatory stimulation to the right TPJ enhanced the subjects’ performance in only tasks involving simultaneous representation of both self and other, while not affecting tasks involving representing the self alone (Santiesteban et al., 2012). Moreover, another study found that TMS to the right TPJ reduced the subjects’ bias in ingroup/outgroup punishment behaviors, suggesting a diminished self-other boundary (Baumgartner et al., 2014). These findings suggest that the TPJ area could serve the function of co-representation of both self and other, rather than self-representation alone (Eddy, 2016; Santiesteban et al., 2012), so as to allow for rapid switching between one’s own perspective and the perspective of another, which is crucial in a successful social engagement.

Putting it all together, the collaboration between the insula and TPJ could serve another important role in associating our internal and external aspects of the self. The TPJ-insula collaboration could constitute the bodily self-consciousness (Park and Blanke, 2019), which could serve as the basis for further co-representation of social information pertaining to both self and other, to enable efficient interactions with the external world.

For a brief summary, the AMPFC, TPJ could serve differential functions in communicating the internal and external environment through collaborations with the insula. Based on the presence or absence of internal sensory signals conveyed through the insula, AMPFC and TPJ could be able to discriminate between self from non-self-signals. Besides these two areas, premotor cortex is also commonly activated during both Exteroceptive- and Mental-self-processing, and showed more activation during Exteroceptive-processing than during Mental-self-processing. The PMC was functionally connected with the

insula (Cauda et al., 2011), and was associated with self-identity (Park and Blanke, 2019). This also supported our hypothesis that PMC was involved in the integration between bodily and external environment information.

4.3. Mental-self-processing, Familiarity and default-mode network

The direct contrast of self > familiarity revealed the crucial difference between the two, which is the insula and PMC. Insula activities are consistently seen in the three levels of self-processing, but not in Familiarity. This indicated that insula activities, or specifically, interoceptive signals, could be the key to discriminating self from familiarity. This discrimination process is likely to be carried out through the collaboration between the insula and other areas such as AMPFC and TPJ. The lack of insula activities in familiarity indicated that when processing familiar signals, we rely mainly on cognitive functions.

Our results showed extensive overlaps between the self, especially Mental-self-processing, and familiarity, in the AMPFC, PCC and left TPJ. These are all key regions of the default-mode network (DMN). DMN has been frequently implicated in higher resting-state brain activities, as well as in spontaneous self-relevant thoughts (Raichle, 2015). The overlap of DMN regions between self and familiarity, indicated that self-processing does indeed share quite some similarities with familiarity-processing, and that the functions of the overlapped regions may not necessarily be self-specific, which is highly consistent with our previous meta-analysis (Qin and Northoff, 2011). However, there is an additional region found in self-processing but not in familiarity, which is the pACC. This finding is also consistent with our previous meta-analysis (Qin and Northoff, 2011). PACC is adjacent to AMPFC, which has an anatomical connection with the insula, and is more robustly involved in Mental-self-processing than in Exteroceptive- and Interoceptive-processing. These findings suggested that pACC could collaborate with AMPFC to form the pathway through which non-bodily external stimuli are incorporated with bodily signals, to be attributed with self-relatedness. Note that in the current results, although pACC is visually seen in the Mental-self-processing results, it did not survive the contrast of Self > Familiarity. One possible explanation for this discrepancy may be that there is also a certain degree of involvement of pACC activities in familiar stimuli, given that self-relevance could sometimes accompany personal familiarity (Cross et al., 2011; Han and Humphreys, 2016; Markus and Kitayama, 1991).

Another region to be mentioned is the TPJ, which showed high involvement in the single analyses of both self and familiarity. On the other hand, contrast analyses showed that the right TPJ was more involved in the self than familiarity, while left TPJ was more involved in familiarity than the self. This is supported by the previous studies, in that right TPJ was involved in self-other distinction (Santiesteban et al., 2015, 2012; Schurz et al., 2013) while left TPJ was involved in reasoning others’ beliefs, intentions and desires (Gallagher et al., 2000). Furthermore, although both Mental-self-processing and Exteroceptive-processing showed consistent activation in the right TPJ in single analyses, in contrast analyses, Exteroceptive-processing showed higher activation probability than Mental-self-processing. This indicated that right TPJ may be more involved in the Exteroceptive-processing than Mental-self-processing. Furthermore, PCC is also shown in both Mental-self-processing and familiarity. PCC may be involved in more general cognitive functions, such as episodic memory retrieval and visual-spatial imagery (Cavanna and Trimble, 2006; Raichle and Raichle, 2001; Wagner et al., 2005). Combining previous findings and the current results (Qin and Northoff, 2011), we suspect that the representation in PCC might be general, i.e. not specific for either self or familiarity. All the results showed regional overlapping between familiarity and self were based on the studies adopting stimulus-induced amplitude change. New methods, such as task-induced change of cortical activity space (Ponce-Alvarez et al., 2015) or trial-to-trial variability (He, 2013), could be helpful to discriminate self from familiarity even in their

common regions.

4.4. The neural model of the self

Based on the above results, we hereby provide neural evidence for a three-level-self model, which highlighted a gradient pattern of self-processing from the internal body to the external environment: Interoceptive-processing (the internal level, integrating interoceptive information), Exteroceptive-processing (the middle level, the link between interoceptive and external environment information), and Mental-self-processing (the external level, collecting self-related non-bodily information from the environment to be integrated with interoceptive information through Exteroceptive-processing). In this model the Interoceptive-processing, with the insula as a core hub, serves as the prerequisite for self-processing, by integrating sensory signals from major survival-based physiological functions, such as cardiovascular and gustatory signals from subcortical areas like the thalamus; the Exteroceptive-processing, with participation of primarily the AMPFC, TPJ and PMC in addition to the insula, links our internal body to the external non-bodily environment. Finally, Mental-self-processing, introduces self-relatedness to external non-bodily stimuli by recruiting even more extensive brain regions such as the VMPFC, DMPFC and PCC, in addition to insula, AMPFC and TPJ. That allows for the integration between internal and external environment information for the self which results in the mental level of self, i.e., mental self. According to this model, Mental-self-processing is grounded upon the incorporation of Interoceptive-processing and Exteroceptive-processing, to attribute self-relatedness to non-bodily external stimuli, and involves extensive neural activities throughout the brain.

Finally, one may wonder whether there is a common mechanism, a “glue” or “common currency” underlying the different levels of self-processing (Northoff et al., 2019). There could be such a mechanism provided by the differential temporo-spatial dynamics within the different levels and their corresponding brain regions (Northoff and Huang, 2017; Northoff and Stanghellini, 2016). If so, one would expect the temporo-spatial features of the resting state in the above mentioned regions and levels to transform in somewhat corresponding temporo-spatial features on the different levels of self-processing. Such temporo-spatial theory of self (Northoff and Huang, 2017; Northoff and Stanghellini, 2016) and self-consciousness (Northoff and Huang, 2017) remains to be tested though.

4.5. Relation between our model and previous concepts of self

It is necessary to discuss the relationship between the current self-model and other important concepts of self. First, William James’ accounts of the “Me” (self as object) and the “I” (self as subject) discussed different natures of the self (James, 1890). Self as object accounts for attributes related to or characterizing the own self, which is ingrained in certain contents. Based on the content, it can be represented as bodily self, autobiographical self, social self, and so on (Northoff, 2016a; Sui and Gu, 2017). It has been proposed that most current paradigms investigating the self reflect self as object (Sui and Gu, 2017), since in these paradigms, self-related stimuli are observed from a reflective/introspective angle (e.g. my face, my personality) and are highly content dependent. Although in everyday life, interoceptive-signals are mostly processed implicitly, which is also the case in some of the Interoceptive-processing tasks, self-processing in tasks like heart-beat counting should also be seen as reflecting self as object. On the other hand, self as subject refers to “being an agent” that experiences itself and the external environment, which is strongly related to self-consciousness, the consciousness of oneself and its’ relationship with surrounding environments (Northoff, 2016a). We consider the self as subject to underlie all self-processing. Therefore, self as subject should be present in Exteroceptive-processing and Mental-self-processing while Interoceptive-processing is implicit, could not be sufficient by

itself for an explicit self as subject.

Our three-level-self model is different from Damasio’ theory about the self which also includes three levels. In his model, the “proto self” generates primordial feelings, which is largely subconscious and provides the grounding for the higher levels of self; the “core self” represents the transient relationship between an individual and the surrounding environment; the “autobiographical self” represents the fundamental sense of self, which requires working and long-term memory (Damasio, 2003b). Like his model, our model is a gradient-based model; however, the different steps are not identical. Within our model, the three levels of self-processing are discriminated according to the relationship between bodily and external environment information. For example, Interoceptive-processing purely involves interoceptive information; Exteroceptive-processing involves external stimuli which have a direct relationship with the organism, i.e. face photo or objects in congruent motion with the body; Mental-self-processing involves non-bodily related external environment stimuli, i.e. a geometric figure. Whether and how our three levels of Interoceptive-, Exteroceptive-, and Mental-self processing correspond to Damasio’s three levels of self, i.e., proto-self, core self, and extended self, remains open to future conceptual investigation.

Self-referential processing and self-related processing are also frequently discussed in the literatures about the self (Christoff et al., 2011; Northoff, 2011). Self-referential processing is the explicit processing where a person is aware that specific contents are related to him/herself, and can lead to self-reference effect (Christoff et al., 2011; Northoff, 2011). For the current model, self-referential processing is not likely to underlie the primarily implicit Interoceptive-processing. For Exteroceptive-processing, it is possible that self-referential processing is involved, as reflected in the “mirror tests”, a common way to detect self-awareness (Anderson, 1984). Finally, Mental-self-processing is involved in the self-reference tasks (i.e., self-related trait judgment), and should mainly involve self-referential processing. As for self-related processing, although Christoff and colleagues suggested that it is interchangeable with self-referential processing (Christoff et al., 2011), we here prefer to make a distinction between these two. Self-related processing may reflect the basic relationship between stimulus and the self, which is independent of awareness and should be implicit and automatic (Northoff, 2016a, 2016b, 2011). All the three levels of self-processing in the current model could involve self-related processing, since interoceptive-processing is mostly automatic and implicit, and implicit processing has also been found to be involved in recognition of one’s own face (Exteroceptive-processing) (Wójcik et al., 2018) and own name (Mental-self-processing) (Qin et al., 2010).

4.6. The three-level-self model and predictive coding

Our self-model is highly compatible with the predictive coding model, which in recent years has gained increasing attention (Friston, 2010) in explaining the neuronal and computational basis of brain functions such as perception (Barrett and Simmons, 2015) and consciousness (Seth et al., 2012). Following the free-energy principle, the predictive coding theory posits that the brain processes information in a hierarchical structure, where in each level, top-down predictions and bottom-up inputs are compared and iterated to minimize (explain away) the discrepancy, i.e., free energy between them (prediction errors) (Friston, 2010). Self-models have also been proposed under the predictive coding account (Apps and Tsakiris, 2014; Limanowski and Blankenburg, 2013). For instance, one paper discussed the predictive coding account of a minimal-self, which is the basic, pre-reflective experience of a bodily-self, to be the emergent result from optimizing predictions about the sensory consequences of events occurring in the environment (Apps and Tsakiris, 2014). The optimization is done by dynamically updating its predictions based on prediction errors between the predicted and actual sensory outcomes (Apps and Tsakiris, 2014).

Our hierarchical self-model fits well with these theories, and provides a more comprehensive framework with strong neural evidence for various levels of self-processing. Combining the predictive coding theory and our three-level model, it is possible that the self-relatedness of any given signal is constituted by being processed through an Interoceptive- to Mental-self-processing hierarchy, by iteratively comparing the signals with the predicted self-relatedness of that signal, and dynamically updating the predictions to minimize the prediction errors, resulting in both bottom-up and top-down modulations. At the interoceptive-processing level, interoceptive signals are compared with the initial predictions of self-relatedness, which, being encoded in the spontaneous activity's spatiotemporal structure (Duncan et al., 2015; Huang et al., 2016; Wolff et al., 2019), are likely to come from early experiences such as before birth (Duncan et al., 2015). Since these early experiences comprise mainly of interoceptive signals (Murphy et al., 2017), these predictions should be highly consistent, or even reliant on interoceptive signals, and should encounter relatively small prediction errors. Then at the second level, Exteroceptive-processing, the incoming external sensory signals bring in larger prediction errors, thus requiring predictions to be updated and expanded, hence recruiting more extended brain areas as we show in our data on the second level. At this level, implicit processing may become explicit. For instance, in an agency task, external sensory signals that have high consistency with prior predictions would be explicitly perceived as self-related (e.g. "my arm is moving"), whereas signals that do not match these predictions (e.g. without simultaneous incoming proprioceptive signals) would be explicitly judged as non-self-related (e.g. "an arm is moving") (Asai, 2017). Furthermore, since predictions at this second more exteroceptive level are built on predictions from the previous interoceptive level, alterations in Interoceptive-processing could, in turn, affect and shape Exteroceptive-processing as it has indeed been supported by previous evidence (Aspell et al., 2013). Similarly, predictions in the Mental-self-processing level have to be updated and expanded from predictions in the previous intero- and exteroceptive levels to account for the ever larger prediction errors brought upon by the growing complexity of non-bodily external information, thus recruiting even more brain regions as we show in our data on the Mental-self-processing level.

The predictive coding account is also consistent with our previous theory that a "rest-self overlap" underlies brain activities, where resting-state activities in primarily the midline regions overlap with self-processing, and could predict different degrees of self-relatedness of incoming signals (Bai et al., 2016; Northoff, 2016b, 2013, 2011). As based on most likely prior and early life events (Duncan et al., 2015) combined with the early interoceptive inputs from the body, these resting-state activities may encode and thus incorporate self-related information (Northoff et al., 2011). That self-related information encoded in resting-state activity (in spatiotemporal terms) (Northoff, 2016a) may be recruited in yielding predictions of self-relatedness of potentially incoming intero- and exteroceptive stimuli. Being exposed to the respective stimuli and processed through the above described iterative hierarchical three-level structure, will then update the predictions of self-relatedness of the respective stimuli through their prediction errors. At the same time, the prediction errors may also modify the underlying spontaneous activity's spatiotemporal structure through stimulus-rest interaction, by means of which the more temporally continuous self-related information is also modified and adapted (Northoff et al., 2010; Schneider et al., 2008). While both lines of processing, predictive coding of self-related stimuli (Apps and Tsakiris, 2014; Seth and Tsakiris, 2018) and enhanced stimulus-rest interaction of self-related information (Schneider et al., 2008) have been demonstrated on separate grounds, future studies are needed to converge and combine them within one experimental design. Our three-level hierarchical model with the continuous spatial extension of different levels may offer an empirically plausible framework for doing so.

4.7. Implications for mental disorders

Schizophrenia is frequently associated with a self-disorder (Parnas and Zandersen, 2018). One of the most prominent symptoms in schizophrenia is hallucinations, and we believe that some hallucinations with schizophrenia are likely to be results of Exteroceptive-processing malfunctioning. Based on our model, Exteroceptive-processing involves brain regions such as the insula, AMPFC and TPJ, and lesions to these regions could result in failure of explicit recognition of the normal internal source of a sensory signal, hence denial/rejection of its self-relevance. There has indeed been evidence showing abnormal activities in these regions in auditory hallucinations (Jardri et al., 2011). Combined with our model, Exteroceptive-processing malfunctioning in the anterior insula and AMPFC or TPJ, might have an abnormal interaction with regions such as the auditory cortex, causing a failure in recognition of the internal source of the sound (someone is speaking to me) (Northoff and Qin, 2011). Besides hallucinations, delusional thoughts are another common symptom in schizophrenic patients. They have been found to have abnormal neural characteristics in the insula, TPJ (Huber et al., 2018), AMPFC (Gao et al., 2015), as well as ACC (Palaniyappan and Liddle, 2012) and dMPFC (Menon et al., 2011). Based on our model, as well as the predictive coding account, delusions could be a result of Mental-self-processing malfunctioning, where slight, but extensive alterations in the predictions of self-relatedness propagated from above regions, result in biased selection of signals that they perceive as self-related (e.g. a random glance from a passer-by could be perceived by the patient as being spied on) and eventually lead to false judgements and delusional thoughts (someone is out to get me).

Autism has also been found to have problems with the self, as reflected in the mirror-self tests (Carmody and Lewis, 2012; Dawson and McKissick, 1984; Lind and Bowler, 2009), and self-referential tasks (Grisdale et al., 2014; Lombardo et al., 2007; Toichi et al., 2002). Autism has frequently been associated with interoceptive abnormality (DuBois et al., 2016; Fiene and Brownlow, 2015; Garfinkel et al., 2016; Quattrocki and Friston, 2014), as well as the insula mal-functioning (Ebisch et al., 2011; Uddin and Menon, 2009). Combined with our model, the abnormal performance in self-recognition, i.e. the mirror test, could be originated from the atypical development in Interoceptive-processing, which could eventually propagate and extend to Mental-self-processing, presented as poor social skills that have been the "hallmarks" of this population.

4.8. Limitations

One issue should be noted. Many of our discussions on interoceptive-processing are based on experiments using heartbeat detection tasks, which has been used as an indicator for interoceptive abilities (Domschke et al., 2010; Schauder et al., 2015). There was evidence supporting this. For instance, subjects' performance in the heartbeat detection/counting tasks has been found to be correlated with other modalities in interoception (Herbert et al., 2012). However, there have also been concerns about this. For instance, it has been found that both immediate and delayed feedback could significantly improve accuracy in heartbeat detection, indicating that subjects' performance could be influenced by non-sensory processes, such as their own beliefs or knowledge (Ring et al., 2015). Another study found that the interoceptive accuracy scores obtained in the heartbeat detection task has a low correlation with actual heartbeats, and could not distinguish subjects with high or low performance, indicating that the ability shown in heartbeat detection tasks might not reflect true interoceptive ability (Zamariola et al., 2018). In the current study, we included various interoceptive functions to address the above issues. However, some experiments about interoceptive-processing were based on heartbeat detection tasks in our meta-analysis. Therefore, conclusions derived from these studies should be viewed with caution. More studies adopting other interoceptive approaches should be performed in the future to

address this issue.

5. Conclusions

The current analyses showed that the insula activation is involved in all three levels of self-processing, indicating that integrated internal sensory signals could be the key underlying self-processing. The schematic model built upon our current data shows a gradient pattern for self-processing, in which the Exteroceptive-processing serves a crucial function in linking the internal and external world. However, exactly how this coordination is performed in our brain requires further experiments. Our study provides a unified theoretical framework for self-processing that can be a roadmap for future studies in other disciplines.

Acknowledgements

Pengmin Qin received support from the National Science Foundation of China (31771249, 31971032), the Major Program of National Social Science Foundation of China (grant number 18ZDA293), Guangdong Basic and Applied Basic Research Foundation (2020A1515011250), Key Realm R&D Program of Guangzhou (202007030005) and Key Realm R&D Program of Guangdong Province (2019B030335001). Georg Northoff acknowledges support from the Canadian Institutes of Health Research (201103MOP-244752-BSB-CECA-179644; 201103CCI-248496-CCI-CECA-179644) and the Michael Smith Foundation (200809EJL-194083-EJL-CECA-179644).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.neubiorev.2020.05.004>.

References

- Anderson, J.R., 1984. The development of self-recognition: a review. *Dev. Psychobiol.* 17 (1), 35–49. <https://doi.org/10.1002/dev.420170104>.
- Apps, M.A.J., Tsakiris, M., 2014. The free-energy self: a predictive coding account of self-recognition. *Neurosci. Biobehav. Rev.* 41, 85–97. <https://doi.org/10.1016/j.neubiorev.2013.01.029>.
- Araujo, H.F., Kaplan, J., Damasio, A., 2013. Cortical midline structures and autobiographical-self processes: an activation-likelihood estimation meta-analysis. *Front. Hum. Neurosci.* 7, 1–10. <https://doi.org/10.3389/fnhum.2013.00548>.
- Asai, T., 2017. Know thy agency in predictive coding: meta-monitoring over forward modeling. *Cogn. Sci.* 51, 82–99. <https://doi.org/10.1016/j.concog.2017.03.001>.
- Aspell, J.E., 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.* 12, 2445–2453. <https://doi.org/10.1177/0956797613498395>.
- Babo-rebello, M., Wolpert, N., Adam, C., Hasboun, D., Tallon-baudry, C., Babo-rebello, M., 2016. Is the cardiac monitoring function related to the self in both the default network and right anterior insula? *Philos. Trans. Biol. Sci.* 371. <https://doi.org/10.1098/rstb.2016.0004>.
- Bai, Y., Nakao, T., Xu, J., Qin, P., Chaves, P., Heinzl, A., Duncan, N., Lane, T., Yen, N.S., Tsai, S.Y., Northoff, G., 2016. Resting state glutamate predicts elevated pre-stimulus alpha during self-relatedness: a combined EEG-MRS study on “rest-self overlap”. *Soc. Neurosci.* 11 (3), 249–263. <https://doi.org/10.1080/17470919.2015.1072582>.
- Barrett, L.F., Simmons, W.K., 2015. Interoceptive predictions in the brain. *Nat. Rev.* 16 (7), 419–429. <https://doi.org/10.1038/nrn3950>.
- Baumgartner, T., Schiller, B., Rieskamp, J., Gianotti, L.R.R., Knoch, D., 2014. Diminishing parochialism in intergroup conflict by disrupting the right temporo-parietal junction. *Soc. Cogn. Affect. Neurosci.* 9 (5), 653–660. <https://doi.org/10.1093/scan/nst023>.
- Beer, J.S., Heerey, E.A., Keltner, D., Scabini, D., Knight, R.T., 2003. The regulatory function of self-conscious emotion: insights from patients with orbitofrontal damage. *J. Pers. Soc. Psychol.* 85 (4), 594–604. <https://doi.org/10.1037/0022-3514.85.4.594>.
- Belk, R.W., 1988. Possessions and the extended self. *J. Consum. Res.* 15 (2), 139. <https://doi.org/10.1086/209154>.
- Blanke, O., 2012. Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* 13 (8), 556–571. <https://doi.org/10.1038/nrn3292>.
- Blanke, O., Slater, M., Serino, A., 2015. Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* 88, 145–166. <https://doi.org/10.1016/j.neuron.2015.09.029>.
- Brener, J., Ring, C., 2016. Towards a psychophysics of interoceptive processes: the measurement of heartbeat detection. *Philos. Trans. R. Soc. B Biol. Sci.* 371. <https://doi.org/10.1098/rstb.2016.0015>.
- Burrows, C.A., Laird, A.R., Uddin, L.Q., 2016. Functional connectivity of brain regions for self- and other-evaluation in children, adolescents and adults with autism. *Dev. Sci.* 19, 564–580. <https://doi.org/10.1111/desc.12400>.
- Buuren, V., Gladwin, T.E., Zandbelt, B.B., Kahn, S., Vink, M., 2010. Reduced functional coupling in the default-mode network during self-referential processing. *Hum. Brain Mapp.* 1127, 1117–1127. <https://doi.org/10.1002/hbm.20920>.
- Cabeza, R., Prince, S.E., Daselaar, S.M., Greenberg, D.L., Budde, M., Dolcos, F., Labar, K.S., Rubin, D.C., 2004. Brain activity during episodic retrieval of autobiographical and laboratory events: an fMRI study using a novel photo paradigm. *J. Cogn. Neurosci.* 16 (5), 1583–1594. <https://doi.org/10.1162/08998929042568578>.
- Camps, G., Veit, R., Mars, M., de Graaf, C., Smeets, P.A., 2018. Just add water: effects of added gastric distention by water on gastric emptying and satiety related brain activity. *Appetite* 127, 195–202. <https://doi.org/10.1016/j.appet.2018.04.023>.
- Carmody, D.P., Lewis, M., 2012. Self representation in children with and without autism spectrum disorders. *Child Psychiatry Hum. Dev.* 43 (2), 227–237. <https://doi.org/10.1007/s10578-011-0261-2>.
- Cauda, F., D’Agata, F., Sacco, K., Duca, S., Geminiani, G., Vercelli, A., 2011. Functional connectivity of the insula in the resting brain. *Neuroimage* 55 (1), 8–23. <https://doi.org/10.1016/j.neuroimage.2010.11.049>.
- Cavanna, A.E., Trimble, M.R., 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. <https://doi.org/10.1093/brain/awl004>.
- Chan, Y.-C., Lavalley, J.P., 2015. Temporo-parietal and fronto-parietal lobe contributions to theory of mind and executive control: an fMRI study of verbal jokes. *Front. Psychol.* 6, 1–13. <https://doi.org/10.3389/fpsyg.2015.01285>.
- Christoff, K., Cosmelli, D., Legrand, D., Thompson, E., 2011. Specifying the self for cognitive neuroscience. *Trends Cogn. Sci.* 15, 104–112. <https://doi.org/10.1016/j.tics.2011.01.001>.
- Committer, G., Pitzalis, S., Galati, G., Patria, F., Pelle, G., Sabatini, U., Castriota-Scanderberg, A., Piccardi, L., Guariglia, C., Pizzamiglio, L., 2007. Neural bases of personal and extrapersonal neglect in humans. *Brain* 130, 431–441. <https://doi.org/10.1093/brain/awl265>.
- Craig, A.D., 2009. How do you feel - now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. <https://doi.org/10.1038/nrn2555>.
- Craig, A.D.B., 2010. The sentient self. *Brain Struct. Funct.* 214 (5–6), 563–577. <https://doi.org/10.1007/s00429-010-0248-y>.
- Critchley, H.D., Wiens, S., Rotshtein, P., Öhman, A., Dolan, R.J., 2004. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195. <https://doi.org/10.1038/nn1176>.
- Cross, S.E., Hardin, E.E., Gercek-Swing, B., 2011. The what, how, why, and where of self-construal. *Pers. Soc. Psychol. Rev.* 15 (2), 142–179. <https://doi.org/10.1177/1088868310373752>.
- Damasio, A., 2003a. Feelings of emotion and the self. *Ann. N. Y. Acad. Sci.* 1001, 253–261. <https://doi.org/10.1196/annals.1279.014>.
- Damasio, A., 2003b. The person within mental self. *Nature* 423, 227. <https://doi.org/10.1038/423227a>.
- Dawson, G., McKissick, F.C., 1984. Self-recognition in autistic children. *J. Autism Dev. Disord.* 14 (4), 383–394. <https://doi.org/10.1007/BF02409829>.
- de Greck, M., Rotte, M., Paus, R., Moritz, D., Thiemann, R., Proesch, U., Bruer, U., Moerth, S., Tempelmann, C., Bogerts, B., Northoff, G., 2008. Is our self based on reward? Self-relatedness recruits neural activity in the reward system. *Neuroimage* 39 (4), 2066–2075. <https://doi.org/10.1016/j.neuroimage.2007.11.006>.
- Denny, B.T., Kober, H., Wager, T.D., Ochsner, K.N., 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), 1742–1752. https://doi.org/10.1162/jocn_a.00233.
- Devue, C., Collette, F., Baetee, E., Degueldre, C., Luxen, A., Maquet, P., Brédart, S., 2007. Here I am: the cortical correlates of visual self-recognition. *Brain Res.* 1143, 169–182. <https://doi.org/10.1016/j.brainres.2007.01.055>.
- Domschke, K., Stevens, S., Pfleiderer, B., Gerlach, A.L., 2010. Interoceptive sensitivity in anxiety and anxiety disorders: an overview and integration of neurobiological findings. *Clin. Psychol. Rev.* 30, 1–11. <https://doi.org/10.1016/j.cpr.2009.08.008>.
- DuBois, D., Ameis, S.H., Lai, M.C., Casanova, M.F., Desarkar, P., 2016. Interoception in autism spectrum disorder: a review. *Int. J. Dev. Neurosci.* 52, 104–111. <https://doi.org/10.1016/j.ijdevneu.2016.05.001>.
- Duncan, N.W., Hayes, D.J., Wiebking, C., Tiret, B., Pietruska, K., Chen, D.Q., Rainville, P., Marjańska, M., Ayad, O., Doyon, J., Hodaie, M., Northoff, G., 2015. Negative childhood experiences alter a prefrontal-insular-motor cortical network in healthy adults: a preliminary multimodal rsfMRI-fMRI-MRS-dMRI study. *Hum. Brain Mapp.* 36 (11), 4622–4637. <https://doi.org/10.1002/hbm.22941>.
- Ebisch, S.J.H., Gallese, V., Willems, R.M., Mantini, D., Groen, W.B., Romani, G.L., Buitelaar, J.K., Bekkering, H., 2011. Altered intrinsic functional connectivity of anterior and posterior insula regions in high-functioning participants with autism spectrum disorder. *Hum. Brain Mapp.* 32, 1013–1028. <https://doi.org/10.1002/hbm.21085>.
- Eddy, C.M., 2016. The junction between self and other? Temporo-parietal dysfunction in neuropsychiatry. *Neuropsychologia* 89, 465–477. <https://doi.org/10.1016/j.neuropsychologia.2016.07.030>.
- Egan, G., Silk, T., Zamarripa, F., Williams, J., Federico, P., Cunningham, R., Carabott, L., Blair-West, J., Shade, R., McKinley, M., Farrell, M., Lancaster, J., Jackson, G., Fox, P., Denton, D., 2003. Neural correlates of the emergence of consciousness of thirst. *Proc. Natl. Acad. Sci.* 100, 15241–15246. <https://doi.org/10.1073/pnas.2136650100>.
- Eickhoff, S.B.S., Laird, A.R.A., Grefkes, C., Wang, L.E., Zilles, K., Fox, P.T., 2009. Coordinate-based activation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty.

- Struct. Funct. 216, 151–157. <https://doi.org/10.1007/s00429-010-0298-1>.
- Stephani, C., Fernandez-Baca Vaca, G., MacLunas, R., Koubeissi, M., Lüders, H.O., 2011. Functional neuroanatomy of the insular lobe. *Brain Struct. Funct.* 216, 137–149. <https://doi.org/10.1007/s00429-010-0296-3>.
- Stern, E.R., Grimaldi, S.J., Muratore, A., Murrrough, J., Leibu, E., Fleysher, L., Goodman, W.K., Burdick, K.E., 2017. Neural correlates of interoception: effects of interoceptive focus and relationship to dimensional measures of body awareness. *Hum. Brain Mapp.* 1–15. <https://doi.org/10.1002/hbm.23811>. 00.
- Stevenson, R.J., Mahmut, M., Rooney, K., 2015. Individual differences in the interoceptive states of hunger, fullness and thirst. *Appetite* 95, 44–57. <https://doi.org/10.1016/j.appet.2015.06.008>.
- Stoeckel, M.C., Esser, R.W., Gamer, M., Büchel, C., von Leupoldt, A., 2015. Brain mechanisms of short-term habituation and sensitization toward dyspnea. *Front. Psychol.* 6, 748. <https://doi.org/10.3389/fpsyg.2015.00748>.
- Sugiura, M., Watanabe, J., Maeda, Y., Matsue, Y., Fukuda, H., Kawashima, R., 2005. Cortical mechanisms of visual self-recognition. *Neuroimage* 24, 143–149. <https://doi.org/10.1016/j.neuroimage.2004.07.063>.
- Sugiura, M., Sassa, Y., Jeong, H., Wakusawa, K., 2011. Self-face recognition in social context. *Hum. Brain Mapp.* 1–11. <https://doi.org/10.1002/hbm.21290>. 000.
- Sui, J., Gu, X., 2017. Self as object: emerging trends in self research. *Trends Neurosci.* 40, 643–653. <https://doi.org/10.1016/j.tins.2017.09.002>.
- Sui, J., Humphreys, G.W., 2015. The integrative self: how self-reference integrates perception and memory. *Trends Cogn. Sci.* 19 (12), 719–728. <https://doi.org/10.1016/j.tics.2015.08.015>.
- Sui, J., Humphreys, G.W., 2016. The ubiquitous self: what the properties of self-bias tell us about the self. *Ann. N. Y. Acad. Sci.* 1396, 222–235. <https://doi.org/10.1111/nyas.13197>.
- Sui, J., Rotshtein, P., Humphreys, G.W., 2013. Coupling social attention to the self forms a network for personal significance. *Proc. Natl. Acad. Sci.* 110 (19), 7607–7612. <https://doi.org/10.1073/pnas.1221862110>.
- Summerfield, J.J., Hassabis, D., Maguire, E.A., 2009. Cortical midline involvement in autobiographical memory. *Neuroimage* 44, 1188–1200. <https://doi.org/10.1016/j.neuroimage.2008.09.033>.
- Suzuki, K., Garfinkel, S.N., Critchley, H.D., Seth, A.K., 2013. Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia* 51, 2909–2917. <https://doi.org/10.1016/j.neuropsychologia.2013.08.014>.
- Tacikowski, P., Brechmann, A., Marchewka, A., Jednoróg, K., Dobrowolny, M., n.d. Is it about the self or the significance? An fMRI study of self-name recognition. *Soc. Neurosci.* 6(1), 98–107. <https://doi.org/10.1080/17470919.2010.490665>.
- Toichi, M., Kamio, Y., Okada, T., Sakihama, M., Youngstrom, E.A., Findling, R.L., Yamamoto, K., 2002. A lack of self-consciousness in autism. *Am. J. Psychiatry* 159 (8), 1422–1424. <https://doi.org/10.1176/appi.ajp.159.8.1422>.
- Tsakiris, M., 2008. Looking for myself: current multisensory input alters self-face recognition. *PLoS One* 3 (12), e4040. <https://doi.org/10.1371/journal.pone.0004040>.
- Tsakiris, M., 2017. The multisensory basis of the self: from body to identity to others. *Q. J. Exp. Psychol.* 70, 597–609. <https://doi.org/10.1080/17470218.2016.1181768>.
- Tsakiris, M., Critchley, H., 2016. Interoception beyond homeostasis: affect, cognition and mental health. *Philos. Trans. R. Soc. B Biol. Sci.* 371. <https://doi.org/10.1098/rstb.2016.0002>.
- Tsakiris, M., Costantini, M., Haggard, P., 2008. The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. *Neuropsychologia* 46, 3014–3018. <https://doi.org/10.1016/j.neuropsychologia.2008.06.004>.
- Turk, D.J., Van Bussel, K., Waiter, G.D., Macrae, C.N., 2011. Mine and me: exploring the neural basis of object ownership. *J. Cogn. Neurosci.* 23 (11), 3657–3668. https://doi.org/10.1162/jocn_a.00042.
- Turkeltaub, P.E., Eden, G.F., Jones, K.M., Zeffiro, T.A., 2002. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* 16, 765–780. <https://doi.org/10.1006/nimg.2002.1131>.
- Uddin, L.Q., Menon, V., 2009. The anterior insula in autism: under-connected and under-examined. *Neurosci. Biobehav. Rev.* 33 (8), 1198–1203. <https://doi.org/10.1016/j.neubiorev.2009.06.002>.
- Uddin, L.Q., Nomi, J.S., Hébert-Seropian, B., Ghaziri, J., Boucher, O., 2017. Structure and function of the human insula. *J. Clin. Neurophysiol.* 34, 300–306. <https://doi.org/10.1097/WNP.0000000000000377>.
- van der Meer, L., Costafreda, S., Aleman, A., David, A.S., 2010. Self-reflection and the brain: a theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neurosci. Biobehav. Rev.* 34, 935–946. <https://doi.org/10.1016/j.neubiorev.2009.12.004>.
- van Veluw, S.J., Chance, S.A., 2014. Differentiating between self and others: an ALE meta-analysis of fMRI studies of self-recognition and theory of mind. *Brain Imaging Behav.* 8, 24–38. <https://doi.org/10.1007/s11682-013-9266-8>.
- Vogeley, K., Fink, G.R., 2003. Neural correlates of the first-person perspective. *Trends Cogn. Sci.* 7, 3–7. [https://doi.org/10.1016/s1364-6613\(02\)00003-7](https://doi.org/10.1016/s1364-6613(02)00003-7).
- Wagner, A.D., Shannon, B.J., Kahn, I., Buckner, R.L., 2005. Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 9 (9), 445–453. <https://doi.org/10.1016/j.tics.2005.07.001>.
- Wang, G.J., Tomasi, D., Backus, W., Wang, R., Telang, F., Geliebter, A., Korner, J., Bauman, A., Fowler, J.S., Thanos, P.K., Volkow, N.D., 2008. Gastric distention activates satiety circuitry in the human brain. *Neuroimage* 39 (4), 1824–1831. <https://doi.org/10.1016/j.neuroimage.2007.11.008>.
- Wang, G., Mao, L., Ma, Y., Yang, X., Cao, J., Liu, X., Wang, J., Wang, X., 2012. Neural representations of close others in collectivistic brains. *Soc. Cogn. Affect. Neurosci.* 7 (2), 222–229. <https://doi.org/10.1093/scan/nsr002>.
- Wiebking, C., de Greck, M., Duncan, N.W., Tempelmann, C., Bajbouj, M., Northoff, G., 2015a. Interoception in insula subregions as a possible state marker for depression - an exploratory fMRI study investigating healthy, depressed and remitted participants. *Front. Behav. Neurosci.* 9, 82. <https://doi.org/10.3389/fnbeh.2015.00082>.
- Wiebking, C., Duncan, N.W., Tiret, B., Hayes, D.J., Doyon, J., Bajbouj, M., Northoff, G., Unit, N., 2015b. GABA in the insula - a predictor of the neural response to Interoceptive Awareness. *Neuroimage* 86, 10–18. <https://doi.org/10.1016/j.neuroimage.2013.04.042>.
- Wójcik, M.J., Nowicka, M.M., Kotlewska, I., Nowicka, A., 2018. Self-face captures, holds, and biases attention. *Front. Psychol.* 8, 2371. <https://doi.org/10.3389/fpsyg.2017.02371>.
- Wolff, A., Di Giovanni, D.A., Gómez-Pilar, J., Nakao, T., Huang, Z., Longtin, A., Northoff, G., 2019. The temporal signature of self: temporal measures of resting-state EEG predict self-consciousness. *Hum. Brain Mapp.* 40 (3), 789–803. <https://doi.org/10.1002/hbm.24412>.
- Zamariola, G., Maurage, P., Luminet, O., Corneille, O., 2018. Interoceptive accuracy scores from the heartbeat counting task are problematic: evidence from simple bivariate correlations. *Biol. Psychol.* 137, 12–17. <https://doi.org/10.1016/j.biopsycho.2018.06.006>.