



Neural representations of the multidimensional self in the cortical midline structures

Chunliang Feng^{a,b,c,1}, Xinyuan Yan^{a,b,1}, Wenhao Huang^{a,b}, Shihui Han^d, Yina Ma^{a,b,*}

^a State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

^b Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing 100875, China

^c College of Information Science and Technology, Beijing Normal University, Beijing 100875, China

^d School of Psychological and Cognitive Sciences, PKU-IDG/McGovern Institute for Brain Research, Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100080, China

ARTICLE INFO

Keywords:

Self
Representational similarity analysis
Cortical midline structures
Identity
Dimension

ABSTRACT

Self-concept consists of self-identity that distinguishes self from other people and knowledge that describes one's own attributes in different dimensions. Because self-concept plays a fundamental role in individuals' social functioning and mental health, behavioral studies have examined cognitive processes of self-identity and self-knowledge extensively. Nevertheless, how different dimensions of the self-concept are organized in multi-voxel neural patterns remains elusive. Here, we address this issue by employing representational similarity analyses of behavioral/theoretical models of multidimensional self-representation and blood oxygen level dependent responses, recorded using functional MRI, to judgments of personality traits, physical attributes and social roles of oneself, a close (one's mother) other, and a distant (celebrity) other. The multivoxel patterns of neural activities in the medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC) distinguished representations of the self from both close and distant others, suggesting a specific neural representation of the self-identity; and distinguished different dimensions of person knowledge of oneself, indicating dimension-sensitive neural representation of the self. Moreover, the pattern of PCC activity is more strongly coupled with dimensions of self-knowledge than self-identity. Our findings suggest that multivoxel neural patterns of the cortical midline structures distinguish not only self from others but also discriminate different dimensions of the self.

1. Introduction

Knowing oneself as a unique entity occurs early in human development and is a prerequisite for an individual's normal social functioning (Rochat, 2003). Self-concept in adults consists of self-identity that distinguishes the self from other people and self-knowledge that describes one's own attributes in different dimensions such as personality traits, physical attributes and social roles (James, 1950). The elaborated construct of self-concept is important for appreciating others' perspectives and feelings and allows successful social interactions (Decety and Sommerville, 2003).

Given the pivotal role of self-concept in human lives, the cognitive and neural mechanisms underlying self-concept have been studied extensively in psychology, neuroscience and psychiatry (Gillihan and

Farah, 2005; Legrand and Ruby, 2009; Northoff et al., 2011; Ma and Han, 2010, 2011). Behavioral research has revealed faster responses to one's own face/name than those of others (Ma and Han, 2009, 2010; Tacikowski and Nowicka, 2010) and better memory of self-related items than those related to others (Klein et al., 1989; Lord, 1980; Ma and Han, 2011). Healthy adults also show more accurate and faster responses to geometric shapes associated with oneself than with familiar/unfamiliar others (Sui et al., 2012, 2013). Functional MRI studies evidenced that the cortical midline structures (CMS) including the medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC) exhibit greater blood oxygen level dependent (BOLD) responses to personality trait judgments of the self than others in the self-referential task (Hu et al., 2016; Kelley et al., 2002; Ma and Han, 2011; 2014a; 2014b; 2014c; Northoff and Bermphol, 2004). However, other fMRI studies have reported that

* Corresponding author. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, 19 Xin Jie Kou Wai Da Jie, Beijing, 100875, China.

E-mail address: yina@bnu.edu.cn (Y. Ma).

¹ C.F. and X.Y. contributed equally to this work.

<https://doi.org/10.1016/j.neuroimage.2018.08.018>

Received 2 February 2018; Received in revised form 5 June 2018; Accepted 10 August 2018

Available online 15 August 2018

1053-8119/© 2018 Elsevier Inc. All rights reserved.

reflection on close others (e.g., one's mother or spouse) also induces activations in the mPFC and PCC (Denny et al., 2012; Han et al., 2016; Krienen et al., 2010; Murray et al., 2012; Romund et al., 2017; Zhu et al., 2007). In addition, previous fMRI studies employing repetition-suppression paradigm have shown that the mPFC fails to discriminate self-referential thought and mentalizing about a similar other (Jenkins et al., 2008). These findings cast doubt over the notion of unique cognitive/neural representations of the self (Gillihan and Farah, 2005; Legrand and Ruby, 2009).

Moreover, behavioral studies have suggested independent cognitive representations of the self in different aspects (Marsh and Craven, 2006; Roberts and Donahue, 1994), which make self-concept unique in terms of its multidimensional structure and elaborate contents (Symons and Johnson, 1997). Despite the behavioral evidence for multidimensional structure and contents of self-concept (Symons and Johnson, 1997), brain imaging studies showed that the mPFC activity commonly subserved different aspects of self-referential thoughts (such as personality traits and social roles) and failed to show evidence for distinct neural underpinnings of different dimensions of self-knowledge (Jenkins and Mitchell, 2011; Ma et al., 2014a; 2014c; Martinelli et al., 2013; Moran et al., 2011; Sui et al., 2012). Thus it remains unclear how different dimensions of self-knowledge (e.g., personality traits, social roles, and physical attributes) are represented in the brain.

Most of the previous fMRI studies of self-concept depended on the mass-univariate analysis of the magnitude of BOLD responses. However, it has been recognized that the representation of stimuli or mental states can be characterized by spatially distributed patterns of neural activity that reflect neural population encoding of external stimuli or internal mental states (Georgopoulos et al., 1986; Haxby et al., 2001; Kamitani and Tong, 2005). The distributed patterns of neural activity cannot be captured by the mass-univariate analyses that rely on the average-level neural activity. Thus the current work examined spatially distributed patterns of neural activity associated with representations of multidimensional self-referential thoughts by conducting multivoxel pattern analysis (MVPA) of BOLD responses during reflection of different aspects (i.e., personality traits, social roles, and physical attributes) of oneself, one's mother (a close other), and a celebrity (a distant other). MVPA has been used to unveil spatially overlapping but distinct activity patterns for different categories of stimuli or mental processes. We adopted the representational similarity analysis (RSA, Kriegeskorte et al., 2006; Nili et al., 2014) — a novel MVPA approach that examines mental representations in terms of neural pattern similarities and assesses the geometry of representation structures of the self and others — of BOLD responses during reflection of personality traits, social roles, and physical attributes of oneself and others.

To this end, we first tested a group of healthy participants in a modified label-shape matching task where participants responded to different geometric shapes that have been associated with different dimensions of the self and others. Behavioral performances in this task provide well-validated measures of cognitive representations of the self and others (Sui et al., 2012, 2013), and allow us to build an independent behavioral dissimilarity matrix (DM) for assessing the similarity between oneself and others and between different dimensions of person knowledge of oneself/others. If different dimensions of the self are constructed more elaborated than those of other people, we would expect that the behavioral responses to shapes associated with different dimensions would be more dissimilar for oneself than others. We then sought to identify spatially distributed patterns of brain activity that differentiate between self and others and between different dimensions of self-knowledge to reveal the neural representations of multidimensional self in the brain. In light of previous findings, it was hypothesized that multi-voxel neural patterns in the CMS would not only differentiate the representation of self from close and distant others but might also distinguish the representation of different dimensions of the self, and result in discriminate neural patterns of self-reflection on different dimensions.

2. Materials & methods

2.1. Participants

Seventy-one college students (47 males; age range, 18–27 years; mean age = 20.86 ± 2.06 years) were recruited in the fMRI experiment as paid volunteers (Supplementary Table 1). Two participants were excluded from data analysis, owing to excessive head movement during scanning, leaving 69 participants (45 males) for fMRI data analysis. An independent sample of 19 participants (9 males; age range, 18–28 years; mean age = 22.74 ± 2.66 years) was recruited for the modified label-shape matching task to construct an independent behavioral DM. All participants were right-handed, reported no history of neurological or psychiatric diagnoses, and had normal or corrected-to-normal vision. Informed consent approved by a local ethics committee was provided prior to the study.

2.2. Stimuli and procedure

A classical self-referential task (Kelley et al., 2002; Ma et al., 2014a) was employed during scanning. Participants were asked to judge whether a given word was appropriate to describe oneself (self-judgment), a gender-matched celebrity (celebrity-judgment), or one's mother (mother-judgment) by pressing one of the two buttons with the index or middle finger. For judgments of each target person, there were 3 categories of words describing the mental (personality traits, such as hard-working, friendly), physical (physical attributes, such as black hair, big eyes) and social (social roles, such as student, American) dimensions of person knowledge, and each category consisted of 80 items adopted from our previous work (Ma et al., 2014a; 2014c; Ma and Han, 2011) (Supplementary Table 2). Half of the 80 items of each category were randomly selected for each participant. In a block fMRI design with 5 functional runs, each of the 9 conditions (Identity (of 3 target persons) \times 3 dimensions of person knowledge) was presented once per run in a single block consisting of 8 items. Each item was presented for 2 s followed by 1 s of central fixation. Two successive blocks were intervened by a 10-s rest while participants viewed a white fixation on a black screen. Different blocks in each scan were presented in a random order.

We modified the label-shape matching task (Sui et al., 2012, 2013) for behavioral assessment of inter-condition similarity between each pair of conditions (e.g., own personality vs. own social attribute, or own personality vs. mother's personality) (see also Stolier and Freeman, 2016 for a similar approach). In the modified label-shape matching task, participants first learned 9 label-shape associations between 9 labels (i.e., 3 identities (self, mother, and celebrity) by 3 dimensions (mental, physical and social attributes)) and 9 geometric shapes (e.g., circle, triangle, etc.). For example, for the social dimension, we asked participants to think of a social role that specifically described oneself (e.g., student), one's mother (e.g., manager), and a celebrity (e.g., athlete), and then to associate each social label with a corresponding geometric shape (association such as student-circle, manager-triangle). Similar procedures were used for the mental and physical dimensions with different shapes.

After the learning procedure, participants were asked to complete a matching task in which they were asked to judge whether the presented shape-label pairing was correctly matched. Participants completed two runs of 72 trials for each dimension. In each trial, a white central fixation was presented for 500 ms and followed by a shape-label pairing of 100 ms. Next, there was a blank interval (800–1200 ms), during which participants had to judge whether the shape was correctly associated with the label as quickly and accurately as possible. A 500-ms feedback (“√” or “×” in white color) was then presented.

Notably, previous studies have shown that participants responded with the fastest responses and highest accuracy to self-related associations in this task (Sui et al., 2012, 2013). Therefore, the self-prioritization effects observed in this task are similar to those using self-related stimuli (e.g., one's own name or face), but the label-shape matching task

additionally rules out possible effects of familiarity of stimuli (Sui et al., 2012). Furthermore, previous evidence has suggested that the behavioral performance in this task was modulated as a function of shared representations, such that (i) behavioral performances for close others are better than those for distant others (Sui et al., 2012); and (ii) performances for ingroup-related labels are better than those for outgroup-related labels, and the behavioral advantage is stronger in participants reporting a greater sense of group identity (Enock et al., 2018; Moradi et al., 2015).

2.3. MRI data acquisition

Functional brain images were acquired using a 3-T Siemens Trio scanner at the Beijing MRI Centre for Brain Research. Blood oxygen level-dependent (BOLD) gradient echo-planar images (EPIs) were obtained using a 12-channel head coil [$64 \times 64 \times 32$ matrix with $3.44 \times 3.44 \times 5.0$ mm spatial resolution, repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle = 90° , field of view (FOV) = 24×24 cm]. A high-resolution T1-weighted structural image ($256 \times 256 \times 144$ matrix with a spatial resolution of $1 \times 1 \times 1.33$ mm, TR = 2530 ms, TE = 3.37 ms, inversion time (TI) = 1100 ms, flip angle = 7°) was subsequently acquired.

2.4. Behavioral data analysis

As in previous studies (Sui et al., 2012, 2013), behavioral performance measures (i.e., accuracy, reaction time) of the label-shape matching task were subject to repeated measures analysis of variances (ANOVAs) with Identity (self, mother, celebrity) and Dimension (mental, physical, social) as within-subjects variables for matched and mismatched pairs, respectively. All multiple comparisons were Bonferroni corrected. Behavioral data of the label-shape matching task were used to construct behavioral DMs that characterize the psychological dissimilarity of two paired conditions for 9 matched associations. The representational dissimilarity of two paired conditions was calculated as 1 minus the Pearson correlation coefficient of behavioral responses in the two paired conditions.

2.5. fMRI data analysis

fMRI data were analyzed using SPM12 (Wellcome Trust Centre for Neuroimaging, London). The functional images were realigned to the first volume and remained unsmoothed and in their native space. We estimated a GLM for each participant with identity (self, mother, and celebrity) and dimension (mental, physical and social) as experimental regressors. The experimental regressors were modeled using a boxcar function across the corresponding 24-s blocks and convolved with the canonical hemodynamic response function. The six movement parameters of the realignment (three translations, three rotations) were also included as nuisance regressors. The resulting GLM was corrected for temporal autocorrelations using a first-order autoregressive model. The estimated beta images corresponding to each condition were then averaged across runs at each voxel and used as activity patterns in the representational similarity analysis implemented with the RSA toolbox (Nili et al., 2014).

We extracted local patterns of neural activity in each voxel in the individual native brain image using a searchlight procedure. For each voxel in the individual native brain image, a sphere with a radius of 15 mm was defined. For each condition, the parameter estimates for each of the N voxels in a given sphere was then extracted to represent an N -dimensional pattern vector. The similarity of neural patterns between each pair of target persons (i.e., self-mother, self-celebrity, mother-celebrity) was then computed by calculating the Pearson correlation coefficient between the values in each pair of patterns. Analogously, the neural-pattern similarity among dimensions (i.e., mental, physical, social) for each target person was also computed. Finally, it is likely that

individuals represent distinct dimensions of different targets in different ways. For instance, participants can easily get access to both their own mental states and physical attributes. However, regarding the judgment of other person, individuals can get access to physical attributes of others but have to infer the mental states of others. Therefore, we performed additional analyses comparing neural pattern similarity between the self and others respectively for each dimension of person-knowledge.

Those correlations coefficients were subject to Fisher z transformation for statistical tests. The resulting z maps were normalized to standard space (resampled to $3 \times 3 \times 3$ mm³ voxels), smoothed (FWHM = 8 mm), and entered into a random-effects analysis. Significant results were identified using a voxel threshold of q (FDR) < 0.05.

Furthermore, we compared the neural-pattern dissimilarity (i.e., the neural DM) with the behavioral or theoretical models (i.e., behavioral or theoretical DMs) in each voxel of the brain using the searchlight procedure (Kriegeskorte et al., 2006). The neural DM was constructed by 1 minus the correlation coefficient between the pattern vectors of each condition pair. Next, the Spearman rank between the neural DM and behavioral or theoretical DMs were computed and assigned to the central voxel of the sphere. As such, the searchlight procedure produced Spearman ρ values on each voxel for each participant, which were then subject to Fisher z transformation for statistical tests. The resulting z maps were then normalized to standard space (resampled to $3 \times 3 \times 3$ mm³ voxels), smoothed (FWHM = 8 mm), and entered into a random effects analysis using one-sample t -tests against zero. Significant results were reported using a threshold of P (FWE) < 0.05 at the voxel level. Finally, the anatomical allocation for all significant coordinates was assessed using the SPM xjview toolbox (<http://www.alivelearn.net/xjview/>) and the automated anatomic labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002).

3. Results

3.1. Distinct neural patterns related to self-identity and multidimensional self-knowledge

We conducted whole-brain paired t -tests to assess multivoxel patterns of neural activity involved in reflection of person identity that differentiated between the self and others. The results corroborated greater neural-pattern similarity between one's mother and celebrity than between oneself and celebrity in the PCC, precuneus, mPFC and anterior cingulate cortex (ACC) (Fig. 1A; voxel-wise q (FDR) < 0.05). Furthermore, the neural-pattern similarity between one's mother and celebrity was larger than that between oneself and one's mother in the PCC/precuneus and mPFC/ACC (Fig. 1B; voxel-wise q (FDR) < 0.05). However, the contrast of oneself/mother neural-pattern similarity vs. oneself/celebrity neural-pattern similarity did not reveal any significant results (even using a voxel threshold of $P < 0.005$, uncorrected). The results demonstrated that the distance in CMS neural representations between self and a close (or a distant) other was larger than that between different others. We next compared the neural-pattern similarity of different target respectively for each dimension, and revealed greater neural similarity between one's mother and celebrity than between oneself and celebrity in the PCC, precuneus, temporal-parietal junction (TPJ), mPFC, and ACC (Supplementary Fig. 1; voxel-wise q (FDR) < 0.05) in the mental dimension. No significant result was found for the social nor physical dimensions (even using a voxel threshold of $P < 0.005$, uncorrected).

To examine whether neural representations of person knowledge in social, mental, and physical dimensions are more dissimilar for oneself than others, we conducted whole-brain paired t -tests of BOLD responses during attribute judgments of the three target persons. The results revealed greater dissimilarity in neural patterns related to the three dimensions of oneself than ones' mother in the PCC (Fig. 1C; voxel-wise q (FDR) < 0.05), and of oneself than a celebrity in the PCC and mPFC (Fig. 1D; voxel-wise q (FDR) < 0.05). Nevertheless, a similar whole-brain analysis comparing brain activity patterns underlying mother- and

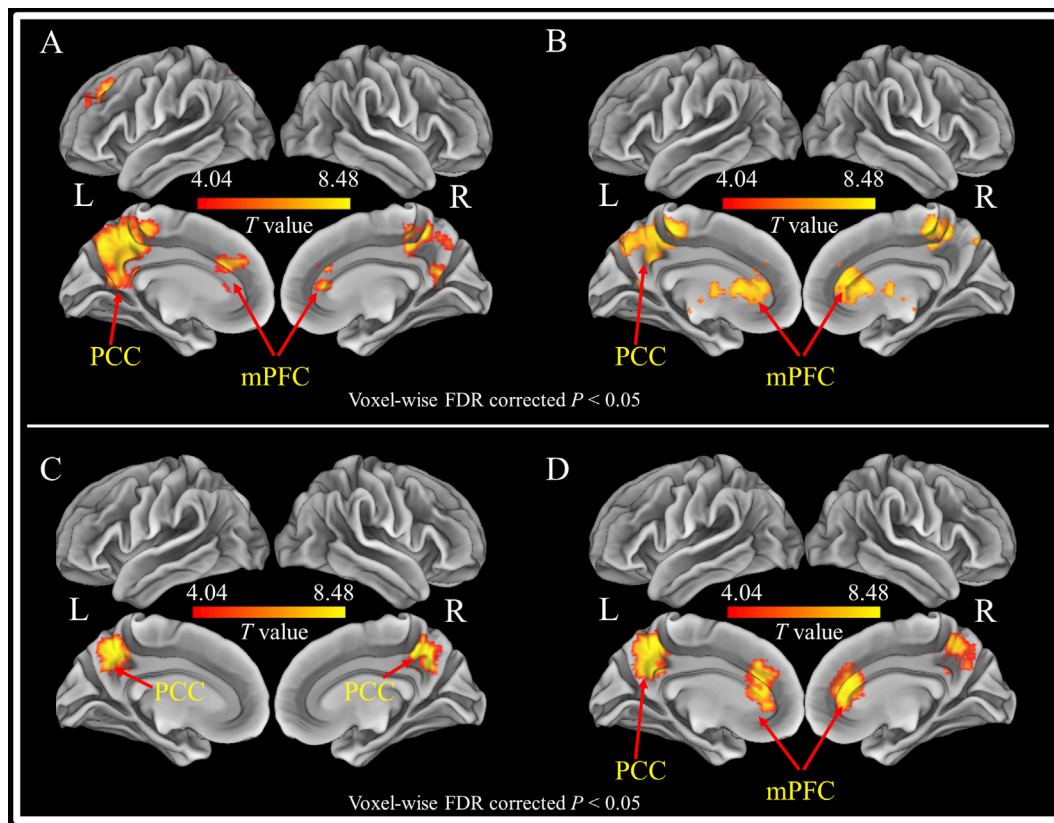


Fig. 1. Pattern similarity between identities and dimensions. **(A)** Neural patterns were more similar between mother and celebrity than between self and celebrity in the following regions: PCC (MNI $x/y/z = 12/-57/12$ mm, cluster size = 77 voxels; maximum $T = 3.71$), precuneus (MNI $x/y/z = -9/-48/48$ mm, cluster size = 985 voxels; maximum $T = 4.93$), mPFC (MNI $x/y/z = 3/36/9$ mm, cluster size = 15 voxels; maximum $T = 3.73$) and anterior cingulate cortex (ACC) (MNI $x/y/z = -12/30/27$ mm, cluster size = 217 voxels; maximum $T = 3.74$). **(B)** Neural patterns were more similar between mother and celebrity than between self and mother in the following regions: PCC/precuneus (MNI $x/y/z = -18/-54/15$ mm, cluster size = 989 voxels; maximum $T = 4.64$) and mPFC/ACC (MNI $x/y/z = 3/33/12$ mm, cluster size = 603 voxels; maximum $T = 4.40$). **(C)** Neural patterns in the PCC (MNI $x/y/z = -3/-57/48$ mm, cluster size = 323 voxels; maximum $T = 6.06$) among dimensions were more similar for mother than self. **(D)** Neural patterns in the PCC (MNI $x/y/z = -3/-54/51$ mm, cluster size = 439 voxels; maximum $T = 6.31$) and mPFC (MNI $x/y/z = 6/33/18$ mm, cluster size = 509 voxels; maximum $T = 6.64$) among dimensions were more similar for celebrity than self. mPFC, medial prefrontal cortex; PCC, posterior cingulate cortex.

celebrity-judgments did not show significant differences in dissimilarity of neural patterns related to the three dimensions of personal knowledge. These results suggest more dissimilar neural representations of person knowledge about the self (relative to others including mother and a celebrity) in the CMS.

3.2. Neural patterns underlying behavioral and theoretical models characterizing multidimensional self

The findings of neural-pattern similarity revealed dissimilar neural patterns of the multidimensional self. Next, we explored the neural patterns corresponding to the behavioral or theoretical models by performing a second-order similarity analysis. We collected behavioral performance in the label-shape matching task. Similar to previous studies (Sui et al., 2012, 2013), the behavioral performance showed a clear boundary between self- and close/distant other-related associations in matched pairs in all dimensions (for details, see Supplementary text and Supplementary Fig. 2). We then constructed a 9×9 behavioral DM that characterized representational dissimilarity of each condition pair (e.g., the distance between oneself and one's mother in personality trait, or, the distance between personality traits and social role of oneself). Dissimilarity was defined as 1 minus the Pearson correlation coefficient of response accuracy within each condition pair (Fig. 2A & Supplementary Fig. 3). As shown in Fig. 2A and B (also see Supplementary Fig. 3), the DM based on response accuracies is characterized by dissimilarity between oneself and others (including one's mother and a celebrity) and

dissimilarity among the three dimensions of oneself (but less so for one's mother and a celebrity). Next, we performed a whole-brain searchlight RSA (Nili et al., 2014) to identify brain regions in which the pairwise similarity of neural patterns of the 9 conditions (3 identities \times 3 dimensions) corresponded to the behavioral DM of condition dissimilarity. The whole-brain searchlight RSA that incorporated the accuracy-based DM revealed that the patterns of neural activity in the mPFC, PCC, and right TPJ corresponded to the behavioral DM (Fig. 2C & Table 1; voxel-wise P (FWE) < 0.05).

To further assess the weight of the brain regions such as the mPFC and PCC in which the activity is engaged in representing person knowledge and person identity, we conducted a whole-brain searchlight RSA that incorporated three theoretical DM models, in which the dissimilarity value was either 0 or 1 (indicating the same or independent representations for each condition pair). The theoretical DMs allowed us to examine the neural patterns that distinguish oneself from others in both person knowledge and person identity, only in person knowledge, or only in person identity. We first conducted a whole-brain searchlight RSA that incorporated the first theoretical DM which considered both unique self-identity and more discriminate dimensional representation for the self and was significantly similar to the behavioral DM (Spearman $\rho = 0.73$, $P < 0.05$). This theoretical DM reliably predicted the patterns of neural activity in the mPFC, PCC, TPJ, and middle temporal gyrus (Fig. 3A & Table 1; voxel-wise P (FWE) < 0.05). We then conducted a whole-brain searchlight RSA that incorporated the second theoretical DM which only distinguishes different dimensions of one's own person knowledge.

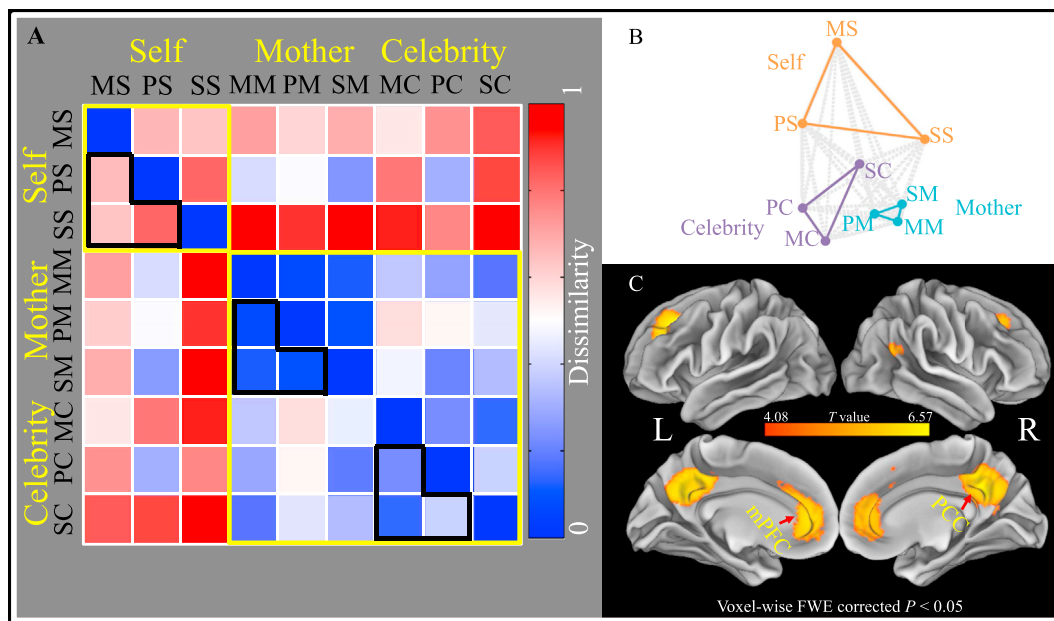


Fig. 2. Behavioral DM and corresponding neural patterns. **(A)** Behavioral DM derived from the label-shape matching task. The behavioral DM is characterized by dissimilarity between oneself and others (including one's mother and a celebrity) and dissimilarity between the three dimensions of person knowledge for oneself (but less so for one's mother and a celebrity). **(B)** Multidimensional scaling of behavioral-pattern similarity. Proximity between points indicates higher behavioral pattern similarity. **(C)** Searchlight results indicating regions predicted by the behavioral DM. The behavioral DM predicted local neural patterns in the mPFC, PCC, and TPJ. MS, mental-self; PS, physical-self; SS, social-self; MM, mental-mother; PM, physical-mother; SM, social-mother; MC, mental-celebrity; PC, physical-celebrity; SC, social-celebrity; mPFC, medial prefrontal cortex; PCC, posterior cingulate cortex; FWE, family-wise error.

This *dimension-sensitive self-representation* model was associated with the activity patterns in the mPFC, PCC, TPJ, middle temporal gyrus (MTG), inferior frontal gyrus (IFG), and dorsolateral prefrontal cortex (dlPFC) (Fig. 3B & Table 1; voxel-wise P (FWE) < 0.05). A whole-brain searchlight RSA that incorporated the third theoretical DM which only distinguishes between oneself from others collapsed dimensions revealed patterns of neural activity in the mPFC and PCC (Fig. 3C & Table 1; voxel-wise P (FWE) < 0.05).

To estimate the relative degree to which the brain regions identified in the above analyses were more strongly associated with discriminate self-knowledge or self-identity, we conducted a paired t -test on the similarity patterns of neural activity revealed in the whole-brain searchlight RSA that incorporated the second and third theoretical DMs. The analysis revealed that activity patterns were more strongly coupled with one's own multidimensional knowledge (vs. self-identity independent of knowledge dimension) in the precuneus/PCC, middle temporal gyrus, and inferior frontal gyrus (voxel-wise P (FWE) < 0.05 , Supplementary Fig. 4 & Supplementary Table 3). However, we did not observe any brain region in which neural patterns were more strongly coupled with one's own person identity (vs. self-knowledge in different dimensions) even with a more liberal threshold ($P < 0.005$, uncorrected).

To examine whether our findings of neural patterns related to one's own person knowledge and person identity were influenced by different calculations of dissimilarity values, we also replicated the whole-brain searchlight RSA using Pearson r or Kendall τ . We found similar results of activity patterns that were coupled with one's own person knowledge and person identity (Fig. 4 & Supplementary Tables 4 and 5).

3.3. Control models of mother and celebrity representations

To reveal brain regions in which neural patterns reflect mother- or celebrity-related representation, we constructed corresponding theoretical models for close and distant others. The searchlight analysis only revealed that the dimension-sensitive mother-representation model predicted the neural-pattern similarity in the middle temporal gyrus and inferior frontal gyrus.

3.4. Univariate activation analysis

Finally, we conducted conventional univariate analyses to examine brain activations indexed by increased BOLD responses associated with different dimensions (mental, physical and social attributes) and identity (self, mother, and celebrity). This analysis of the identity effect revealed stronger activation in the mPFC (among other regions) associated with oneself than one's mother and stronger mPFC activation to one's mother than to a celebrity (Fig. 5 & Supplementary Table 6). The effect of dimensions of personal knowledge was associated with stronger activation in the mPFC and PCC (among other regions) associated with the social dimension than with the mental or physical dimension. Furthermore, the mPFC showed stronger activation to the mental dimension than to the physical dimension, whereas the PCC showed stronger activation to the physical dimension than to the mental dimension (Fig. 5 & Supplementary Table 7).

4. Discussion

The current work investigated neural representations of the multidimensional self by examining the spatially distributed patterns of neural activity to different dimensions of the self and others. Our results revealed multivoxel neural patterns in the CMS characterizing self-identity and multidimensional self-representations. Specifically, the self-identity (especially in the mental dimension) was evident in the multivariate patterns of activity in the mPFC, PCC, and precuneus, which significantly differentiate the self and close/distant others. The results indicated that self-identity was associated with the unique activity patterns in the CMS, although results from univariate analysis indicated that increased average CMS activation during reflection of both the self and close others than distant others (Denny et al., 2012; Han et al., 2016; Kelley et al., 2002; Ma and Han, 2011; Moran et al., 2006; Murray et al., 2012; Zhu et al., 2007). In addition, our RSA results revealed the finer-grained dimension structures of self-representation, as the multivariate patterns of activity in the mPFC, PCC, TPJ, MTG, IFG, and dlPFC differentiate between the three (social, mental, and physical) dimensions

Table 1
Brain regions from searchlight results for the behavioral and theoretical models characterizing unique representations of the self.

Model	Region	L/ R	x/y/z (MNI)	t- value	cluster size
Behavioral DM	Posterior cingulate gyrus	R	15/- 51/27	6.82	1290
	Precuneus	L	-6/- 54/33	5.86	
	Medial frontal gyrus	L	-18/ 30/33	6.26	1460
	Medial frontal gyrus	R	18/27/ 39	4.91	
	Medial frontal gyrus	L	-6/ 45/6	5.25	
	Temporal parietal junction	R	57/- 60/24	4.30	50
	Supramarginal gyrus	R	51/- 54/21	4.28	
	Self-uniqueness DM	Posterior cingulate gyrus	L	-9/- 45/39	8.48
Medial frontal gyrus		L	-3/ 39/-6	7.26	1950
Temporal parietal junction		L	-39/- 60/21	5.64	232
Middle temporal gyrus		L	-60/- 3/-24	4.57	18
Middle temporal gyrus		L	-66/- 6/-18	4.36	
Identity-sensitive self-representation DM		Posterior cingulate gyrus	R	3/-45/ 42	5.81
	Medial frontal gyrus	L	-3/ 39/-6	5.77	836
	Medial frontal gyrus	L	-21/ 33/30	4.84	
	Dimension-sensitive self-representation DM	Posterior cingulate gyrus	L	-9/- 45/39	9.59
Medial frontal gyrus		L	-3/- 42/-6	4.71	342
Middle frontal gyrus		L	-21/ 27/39	5.98	747
Middle frontal gyrus		R	30/24/ 36	4.62	95
Inferior frontal gyrus		R	39/42/ 3	4.27	14
Middle temporal gyrus		L	-63/- 9/-21	4.46	14

DM, dissimilarity matrix; L, left; R, right. Voxel-wise $P(\text{FWE}) < 0.05$.

of person knowledge of oneself but less so for others. The results were replicated using behavioral and theoretical models and different computational approaches (Pearson/Spearman/Kendall correlation coefficients).

The findings have important implications for understanding neural representations of one's own identity and person knowledge. First, we demonstrated that the multivariate patterns of neural activity in the CMS play a key role in coding self-identity. Self-judgments were different from judgments of close/distant others in the multivariate patterns of neural activity in the CMS, whereas close and distant others evoked similar distributed activity patterns. Additional analyses revealed that the unique neural representations of the self-identity were evident in the mental dimension but not in physical and social dimensions. These findings may not be surprising given that people are able to perceive the physical and social attributes of both the self and others, thus the underlying neural substrates may be engaged in a similar manner in these observable dimensions. In contrast, people can get access to their own mental states, but the mental states of other need to be understood via inference/mentalizing. As such, one would expect that differences in neural patterns between self and others would be larger in the mental

dimension than the physical and social dimensions.

Second, we showed that the multivariate patterns of activity in the mPFC, PCC, TPJ, MTG, IFG, and dlPFC were associated with representations of the finer-grained structures of person knowledge of oneself, whereas person knowledge about close and distant others cannot be distinguished by neural patterns. The results suggest that the mPFC and PCC are involved in coding both the identity-sensitive and dimension-sensitive self-representation. However, distinct representations of multidimensional self were also manifested in the multivariate patterns of activity in a neural circuit consisting of brain regions related to autobiographical memory (e.g., the MTG and IFG) (Svoboda et al., 2006), mental states inference (e.g., the TPJ) (Saxe and Kanwisher, 2003), and self-regulation (e.g., the dlPFC) (Smith and Jonides, 1999). The distributed patterns of activity in these brain regions indicate that, rather than being a monolithic construct, self-concept is mediated by complex patterns of activity in a large-scale neural network. These findings provide a neural basis for mental representations of self-knowledge in different dimensions (James, 1950) and for the superior elaborative properties of self-concept (Symons and Johnson, 1997).

Our findings also help to understand the different functional roles of the mPFC and PCC in mediating self-representation. Although both the mPFC and PCC were activated during self-related processing (D'argembeau et al., 2005; Herold et al., 2016; Kurczek et al., 2015; Moran et al., 2006), the mPFC and PCC may contribute to distinct neuropsychological processes underlying self-representations (Northoff and Bermpohl, 2004). Specifically, the mPFC is mainly involved in encoding the self-relevance of stimuli (Northoff et al., 2006; van der Meer et al., 2010) and the PCC integrates self-referential stimuli for more elaborate representations (Northoff and Bermpohl, 2004). Our results of multivoxel patterns of neural activity suggest that both the mPFC and PCC are engaged in representations of self-identity and self-knowledge in different dimensions. On the other hand, neural patterns in the PCC were more strongly associated with representations of self-knowledge than self-identity. These findings suggest that the PCC is involved in the elaboration of dimension sensitive self-knowledge in addition to self-other identity distinction. Consistent with this idea, the PCC has been supposed to play a pivotal role in elaborating self-representations by providing associated autobiographical information (Martinelli et al., 2013; Northoff and Bermpohl, 2004; Svoboda et al., 2006; van der Meer et al., 2010). The distinct patterns of the PCC activity associated with different dimensions of self-knowledge may result from specific information that is retrieved and integrated for each dimension.

Notably, our findings indicated that distinct neural representations of social, mental and physical attributes were evident in the brain regions such as the mPFC and TPJ for the self but less so for close or distant others. This is different from previous findings that neural representations of another person's mental states are also organized according to specific dimensions in brain regions mediating mentalizing (e.g., dorsal mPFC, TPJ) (Saxe and Powell, 2006; Tamir et al., 2016). In previous studies, participants were asked to infer another person's mental states in relatively elaborate contexts. Under such conditions, people might refer to their own thoughts and feelings, which are dimension-sensitive, as a basis for inferring those of others (Jenkins et al., 2008; Nickerson, 1999). Furthermore, vivid contexts corresponding to each dimension might also facilitate discrimination among dimensions of others. In the current study, however, participants were required only to make judgments on abstract knowledge about the self and others. Additionally, it is possible that mental representations of the self and others are organized along different sets of dimensions. Indeed, recent evidence has indicated that representations of others' mental states are intrinsically organized according to the dimensions of rationality, social impact, valence and human mind (Tamir et al., 2016), which are different from the dimensions (social, mental and physical) manipulated in the current study. In this regard, further work should clarify whether the mental states and attributes of the self and others in different dimensions are represented in distinct patterns of brain activity.

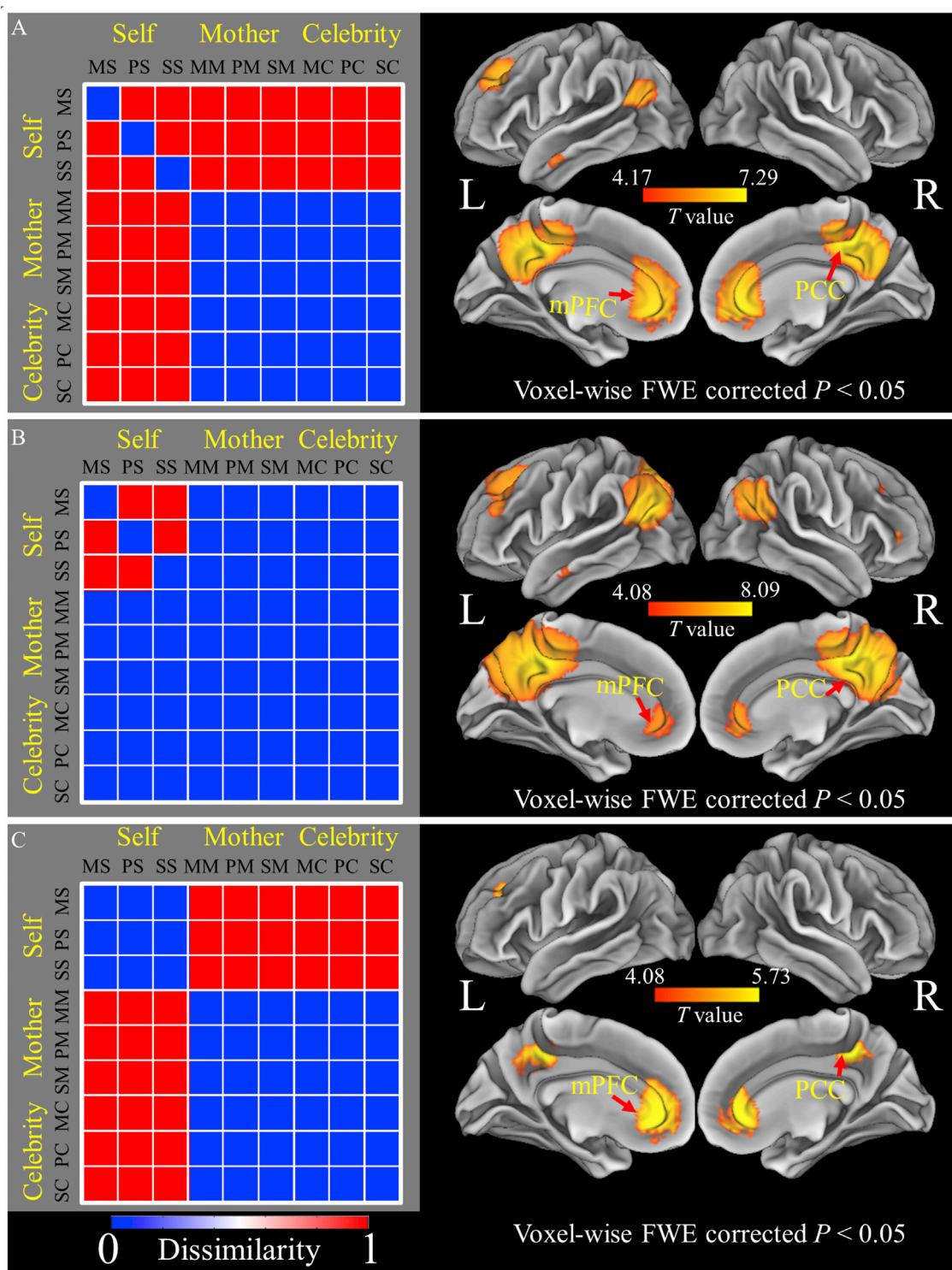


Fig. 3. Theoretical models and corresponding neural patterns. (A) The self-uniqueness theoretical DM and searchlight results of regions predicted by the DM. The self-uniqueness theoretical model predicted local neural patterns in the mPFC, PCC, TPJ, and middle temporal gyrus. (B) The dimension-sensitive self-representation theoretical DM and searchlight results of regions predicted by the DM. The dimension-sensitive self-representation theoretical model predicted local neural patterns in the mPFC, PCC, TPJ, middle temporal gyrus, inferior frontal gyrus, and dIPFC. (C) The identity-sensitive self-representation theoretical DM and searchlight results of regions predicted by the DM. The identity-sensitive self-representation theoretical model predicted local neural patterns in the mPFC and PCC.

Together, our brain imaging findings showed evidence for neural representations of self-identity and self-knowledge in different dimensions in the multivoxel patterns of neural activity across multiple brain regions, which are associated with both behavioral and theoretical models characterizing identity-sensitive and/or dimension-

sensitive representations of the self-concept. Our results shed new light on the nature of neural representations of self-identity and different dimensions of self-knowledge. These findings have important implications for understanding the neural underpinnings of self-concept in the human brain.

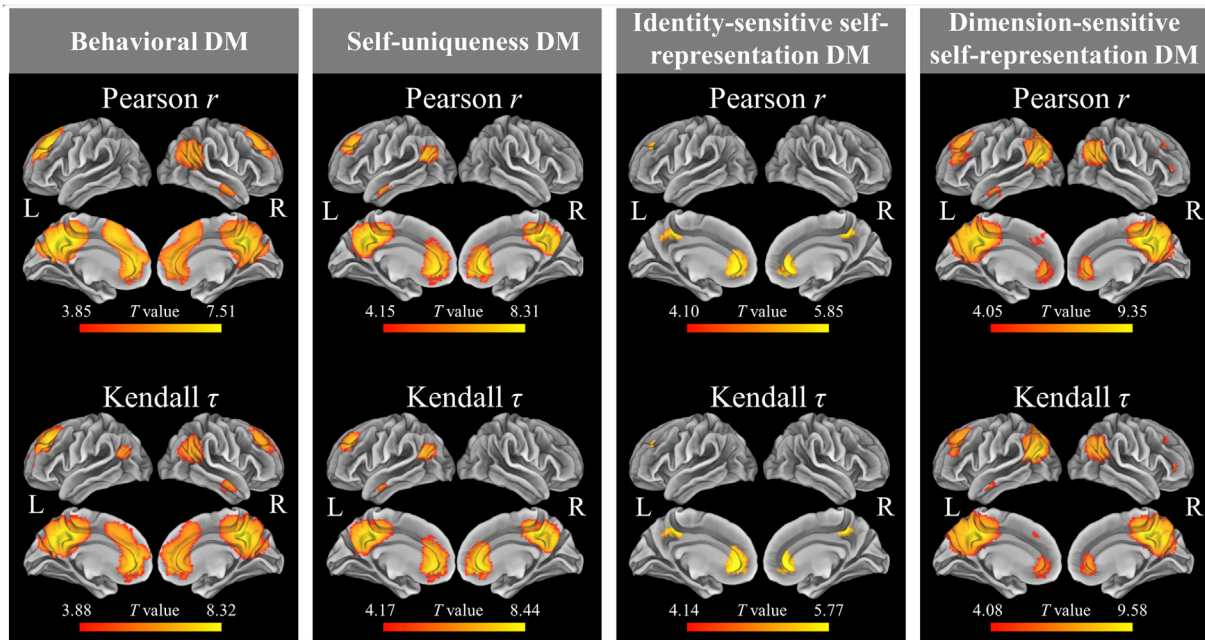


Fig. 4. Searchlight results of self-related models using Pearson r and Kendall τ . The validation analyses with Pearson r or Kendall τ showed similar results of activity patterns in the mPFC and PCC reflecting identity specific dimension specific representations of the self. L, left; R, right; DM, dissimilarity matrix. All maps were thresholded with voxel-wise P (FWE) < 0.05 .

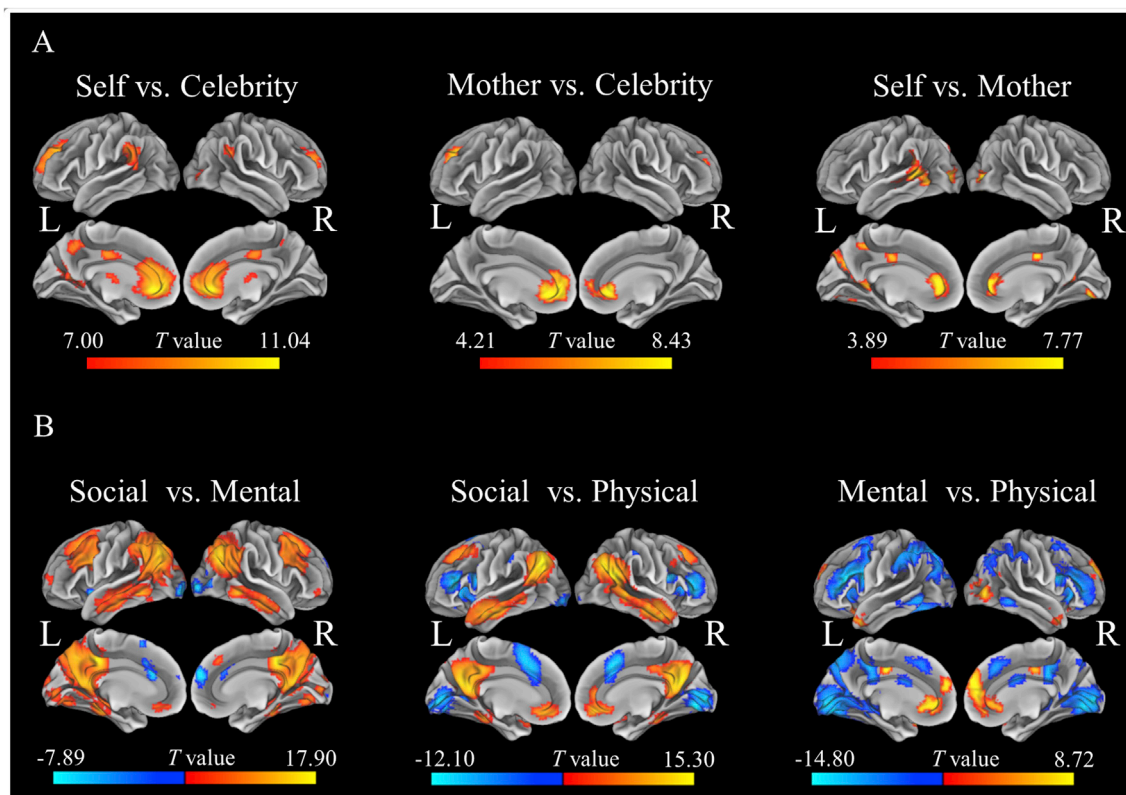


Fig. 5. Univariate activation results regarding the differences between each pair of identities across dimensions (A) and between each pair of dimensions across identities (B). The effect of Identity revealed stronger activations in the mPFC, supramarginal gyrus, precuneus/PCC, middle occipital gyrus, cuneus, middle temporal gyrus, and cerebellum to the self than to the mother and mPFC and superior frontal gyrus to the mother than to a celebrity; no significant activation was found in the reverse contrast. The results regarding the effect of dimension revealed stronger activation in the mPFC, PCC, middle frontal gyrus, and middle temporal gyrus to the social dimension than to the mental or physical dimension. Furthermore, the mPFC, superior frontal gyrus, inferior frontal gyrus, and cerebellum showed stronger activation to the mental dimension than to the physical dimension, whereas the PCC, middle frontal gyrus, hippocampus, and thalamus showed stronger activation to the physical dimension than to the mental dimension. All maps were thresholded with voxel-wise P (FWE) < 0.05 .

Conflicts of interest

No conflict of interest was declared.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Projects 31722026; 31771204; 91632118; 31421003; 31661143039); Open Research Fund of the State Key Laboratory of Cognitive Neuroscience, Beijing Normal University; the Fundamental Research Funds for the Central Universities (2016NT05; 2017XTCX04; 2018EYT04); Beijing Municipal Science and Technology Commission (Z151100003915122); startup funding from the State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, Beijing Normal University; the National Postdoctoral Program for Innovative Talents (BX201600019), and the China Postdoctoral Science Foundation (2017M610055).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.neuroimage.2018.08.018>.

References

- D'argembeau, A., Collette, F., Van der Linden, M., Laureys, S., Del Fiore, G., Degueldre, C., Luxen, A., Salmon, E., 2005. Self-referential reflective activity and its relationship with rest: a PET study. *Neuroimage* 25, 616–624.
- Decety, J., Sommerville, J.A., 2003. Shared representations between self and other: a social cognitive neuroscience view. *Trends Cognit. Sci.* 7, 527–533.
- 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. Cognit. Neurosci.* 24, 1742–1752.
- Enock, F., Sui, J., Hewstone, M., Humphreys, G.W., 2018. Self and team prioritisation effects in perceptual matching: evidence for a shared representation. *Acta Psychol.* 182, 107–118.
- Georgopoulos, A.P., Schwartz, A.B., Kettner, R.E., 1986. Neuronal population coding of movement direction. *Science* 233, 1416–1419.
- Gillihan, S.J., Farah, M.J., 2005. Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychol. Bull.* 131, 76–97.
- Han, S., Ma, Y., Wang, G., 2016. Shared neural representations of self and conjugal family members in Chinese brain. *Culture and Brain* 4, 72–86.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Herold, D., Spengler, S., Sajonz, B., Usnich, T., Bermpohl, F., 2016. Common and distinct networks for self-referential and social stimulus processing in the human brain. *Brain Struct. Funct.* 221, 3475–3485.
- Hu, C., Di, X., Eickhoff, S.B., Zhang, M., Peng, K., Guo, H., Sui, J., 2016. Distinct and common aspects of physical and psychological self-representation in the brain: a meta-analysis of self-bias in facial and self-referential judgements. *Neurosci. Biobehav. Rev.* 61, 197–207.
- James, W., 1950. *The Principles of Psychology: in 2 Volumes*. Dover Publications, New York.
- Jenkins, A.C., Macrae, C.N., Mitchell, J.P., 2008. Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proc. Natl. Acad. Sci. Unit. States Am.* 105, 4507–4512.
- Jenkins, A.C., Mitchell, J.P., 2011. Medial prefrontal cortex subserves diverse forms of self-reflection. *Soc. Neurosci.* 6, 211–218.
- Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685.
- Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F., 2002. Finding the self? An event-related fMRI study. *J. Cognit. Neurosci.* 14, 785–794.
- Klein, S.B., Loftus, J., Burton, H.A., 1989. Two self-reference effects: the importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-referent encoding. *J. Pers. Soc. Psychol.* 56, 853–865.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3863–3868.
- Krienen, F.M., Tu, P.-C., Buckner, R.L., 2010. Clan mentality: evidence that the medial prefrontal cortex responds to close others. *J. Neurosci.* 30, 13906–13915.
- Kurzcek, J., Wechsler, E., Ahuja, S., Jensen, U., Cohen, N.J., Tranel, D., Duff, M., 2015. Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia* 73, 116–126.
- Legrand, D., Ruby, P., 2009. What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychol. Rev.* 116, 252–282.
- Lord, C.G., 1980. Schemas and images as memory aids: two modes of processing social information. *J. Pers. Soc. Psychol.* 38, 257.
- Ma, Y., Bang, D., Wang, C., Allen, M., Frith, C., Roepstorff, A., Han, S., 2014a. Sociocultural patterning of neural activity during self-reflection. *Soc. Cognit. Affect Neurosci.* 9, 73–80.
- Ma, Y., Han, S., 2009. Self-face advantage is modulated by social threat—Boss effect on self-face recognition. *J. Exp. Soc. Psychol.* 45, 1048–1051.
- Ma, Y., Han, S., 2010. Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 619–633.
- Ma, Y., Han, S., 2011. Neural representation of self-concept in sighted and congenitally blind adults. *Brain* 134, 235–246.
- Ma, Y., Li, B., Wang, C., Shi, Z., Sun, Y., Sheng, F., Zhang, Y., Zhang, W., Rao, Y., Han, S., 2014b. 5-HTTLPR polymorphism modulates neural mechanisms of negative self-reflection. *Cerebr. Cortex* 24, 2421–2429.
- Ma, Y., Wang, C., Li, B., Zhang, W., Rao, Y., Han, S., 2014c. Does self-construal predict activity in the social brain network? a genetic moderation effect. *Soc. Cognit. Affect Neurosci.* 9, 1360–1367.
- Marsh, H.W., Craven, R.G., 2006. Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1, 133–163.
- Martinelli, P., Sperduti, M., Piolino, P., 2013. Neural substrates of the self-memory system: new insights from a meta-analysis. *Hum. Brain Mapp.* 34, 1515–1529.
- Moradi, Z., Sui, J., Hewstone, M., Humphreys, G.W., 2015. In-group modulation of perceptual matching. *Psychonomic Bull. Rev.* 22, 1255–1277.
- Moran, J., Macrae, C., Heatherton, T.F., Wyland, C., Kelley, W.M., 2006. Neuroanatomical evidence for distinct cognitive and affective components of self. *J. Cognit. Neurosci.* 18, 1586–1594.
- Moran, J.M., Lee, S.M., Gabrieli, J.D., 2011. Dissociable neural systems supporting knowledge about human character and appearance in ourselves and others. *J. Cogn. Neurosci.* 23, 2222–2230.
- Murray, R.J., Schaer, M., Debbané, M., 2012. Degrees of separation: a quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self-and other-reflection. *Neurosci. Biobehav. Rev.* 36, 1043–1059.
- Nickerson, R.S., 1999. How we know—and sometimes misjudge—what others know: imputing one's own knowledge to others. *Psychol. Bull.* 125, 737.
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., Kriegeskorte, N., 2014. A toolbox for representational similarity analysis. *PLoS Comput. Biol.* 10, e1003553.
- Northoff, G., Bermpohl, F., 2004. Cortical midline structures and the self. *Trends Cognit. Sci.* 8, 102–107.
- Northoff, G., Heinzl, A., De Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J., 2006. Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage* 31, 440–457.
- Northoff, G., Qin, P., Feinberg, T.E., 2011. Brain imaging of the self—conceptual, anatomical and methodological issues. *Conscious. Cognit.* 20, 52–63.
- Roberts, B.W., Donahue, E.M., 1994. One personality, multiple selves: integrating personality and social roles. *J. Personality* 62, 199–218.
- Rochat, P., 2003. Five levels of self-awareness as they unfold early in life. *Conscious. Cognit.* 12, 717–731.
- Romund, L., Golde, S., Lorenz, R.C., Raufelder, D., Pelz, P., Gleich, T., Heinz, A., Beck, A., 2017. Neural correlates of the self-concept in adolescence—a focus on the significance of friends. *Hum. Brain Mapp.* 38, 987–996.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage* 19, 1835–1842.
- Saxe, R., Powell, L.J., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* 17, 692–699.
- Smith, E.E., Jonides, J., 1999. Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661.
- Stolier, R.M., Freeman, J.B., 2016. Neural pattern similarity reveals the inherent intersection of social categories. *Nat. Neurosci.* 19, 795–797.
- Sui, J., He, X., Humphreys, G.W., 2012. Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1105–1117.
- Sui, J., Rotshtein, P., Humphreys, G.W., 2013. Coupling social attention to the self forms a network for personal significance. *Proc. Natl. Acad. Sci. Unit. States Am.* 110, 7607–7612.
- Svoboda, E., McKinnon, M.C., Levine, B., 2006. The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia* 44, 2189–2208.
- Symons, C.S., Johnson, B.T., 1997. The self-reference effect in memory: a meta-analysis. *Psychol. Bull.* 121, 371.
- Tacikowski, P., Nowicka, A., 2010. Allocation of attention to self-name and self-face: an ERP study. *Biol. Psychol.* 84, 318–324.
- Tamir, D.I., Thornton, M.A., Contreras, J.M., Mitchell, J.P., 2016. Neural evidence that three dimensions organize mental state representation: rationality, social impact, and valence. *Proc. Natl. Acad. Sci. Unit. States Am.* 113, 194–199.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.
- van 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.
- Zhu, Y., Zhang, L., Fan, J., Han, S., 2007. Neural basis of cultural influence on self-representation. *Neuroimage* 34, 1310–1316.