OXFORD

Multivariable pattern classification differentiates relational self-esteem from personal self-esteem

Jiwen Li,[1,*] Mei Zeng,[1,*] Mingyan Liu,[1] Xiaolin Zhao,[1] Weiyu Hu,[1] Chong Wang,[2] Chijun Deng,[2] Rong Li,[2] Huafu Chen,[2] and Juan Yang [1]

[1]Key Laboratory of Cognition and Personality, Ministry of Education, Faculty of Psychology, Southwest University, Chongqing 400715, China, and [2]School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 611731, China

*Joint first authorship.

Correspondence should be addressed to Juan Yang, Key Laboratory of Cognition and Personality, Ministry of Education, Faculty of Psychology, Southwest University, Tiansheng Road No.2, Beibei District, Chongqing 400715, China.
E-mail: valleyqq@swu.edu.cn.

Abstract

Relational self-esteem (RSE) refers to one's sense of self-worth based on the relationship with significant others, such as family and best friends. Although previous neuroimaging research has investigated the neural processes of RSE, it is less clear how RSE is represented in multivariable neural patterns. Being able to identify a stable RSE signature could contribute to knowledge about relational self-worth. Here, using multivariate pattern classification to differentiate RSE from personal self-esteem (PSE), which pertains to self-worth derived from personal attributes, we obtained a stable diagnostic signature of RSE relative to PSE. We found that multivariable neural activities in the superior/middle temporal gyrus, precuneus, posterior cingulate cortex, dorsal medial Prefrontal Cortex (dmPFC) and temporo-parietal junction were responsible for diagnosis of RSE, suggesting that the evaluation of RSE involves the retrieval of relational episodic memory, perspective-taking and value calculation. Further, these diagnostic neural signatures were able to sensitively decode neural activities related to RSE in another independent test sample, indicating the reliability of the brain state represented. By providing a reliable multivariate brain pattern for RSE relative to PSE, our results informed more cognitively prominent processing of RSE than that of PSE and enriched our knowledge about how relational self-worth is generated in the brain.

Key words: relational self-esteem (RSE); personal self-esteem (PSE); cognitive demanding; multivariable pattern analysis (MVPA)

Introduction

Relational self-esteem (RSE) refers to one's sense of self-worth based on the relationship with significant others, such as family and friends (Chen *et al.*, 2011; Du *et al.*, 2012). Considering the increasing evidence that the personality system has its corresponding neural mechanisms in the brain (Kosslyn, 1999; Ochsner and Lieberman, 2001), a recent neuroimaging study uncovered the neural substrate of RSE (Li *et al.*, 2019). Nevertheless, this research depended on the traditional univariate analyses of the magnitude of the blood oxygenation level-dependent (BOLD) response, and it remains unclear how RSE is organized in spatially distributed neural patterns. To this end, the current study adopted multivariate pattern

classification to investigate the widely distributed brain signature of RSE. In addition, to eliminate some unrelated processing related to RSE, such as social desirability and self-presentation motivation, the brain signature of RSE was compared with that of personal self-esteem (PSE) (Greenwald and Satow, 1970; Baumeister et al., 1989).

RSE vs PSE

Self-esteem was defined as a person's sense of self-worth (Crocker and Wolfe, 2001) and can be rooted in either personal characteristics or one's relational identity (Tajfel and Turner, 1986). PSE pertains to self-worth derived from personal attributes, such as abilities and talents (e.g. 'I am intelligent' or 'I am incompetent') (Rosenberg, 1965), while the RSE is derived from one's relationship with significant others (e.g. 'I am a worthy member of my circle of friends' or 'I am proud of my family') (Du et al., 2012, 2013). As an important predictor of subjective well-being, high RSE is associated with greater life satisfaction, positive affect, meaning in life and subjective vitality (Wagner, 2009; Du et al., 2017). Moreover, low RSE or lack of RSE can lead to a series of mental health problems, such as depression and anxiety (Du et al., 2013).

Essentially, self-esteem reflects an attitude toward self, which is featured by cognitive and affective processes (Rosenberg et al., 1995; Moran et al., 2006). RSE pertains to self-worth derived from one's relationship with significant others. Compared with PSE, which is derived from one's attributes, RSE's evaluative domains depend on the relatively concrete relational situation, such as family and friends. Moreover, RSE highlights how self-worth is generated both from introspection and from a third perspective (Tajfel and Turner, 1986; Du et al., 2012), whereas PSE only emphasizes the processing of introspection. It is more difficult to conceptualize beliefs and attitudes of others, whose thoughts are not accessible to us, than to conceptualize one's own personal attitudes and beliefs (Murray et al., 2015). Thus, the evaluation of RSE is more cognitively demanding than that of PSE (Rosenberg, 1995; Baumeister et al., 2003). Relatively, affective processing, the representation of PSE, including stable characteristics which could generalize across situations, was relatively prominent when personal self-worth was evaluated, consistent with earlier research which emphasized that PSE is an affectively laden self-evaluation (Leary, 2000; Brown et al., 2001).

Brain patterns for self-esteem

Most previous neuroimaging studies invariably focused on the neural correlates of PSE and found that the brain regions related to PSE were widely distributed in the cortical midline structures (CMSs), which include the medial prefrontal cortex (mPFC), the orbitofrontal cortex, the dorsal anterior cingulate gyrus, the posterior cingulate cortex (PCC) and reward networks in the striatum and caudate (Eisenberger et al., 2011; Yang et al., 2012, 2016; Chavez and Heatherton, 2014; Izuma et al., 2018). These regions suggested that evaluative processing of PSE is related to the retrieval of episodic memory and value assignments (Frewen et al., 2013; Pan et al., 2015). To our knowledge, there is only one neuroimaging study designed by authors to investigate the neural correlates of RSE, and those results found similar activation as in PSE, including activation in the mPFC, the PCC and the precuneus (Li et al., 2019). Meanwhile, the theory of mind network located at the bilateral temporoparietal junction (TPJ), which associates with the evaluation of the value of significant others, was only recruited in the processing of RSE (Li et al., 2019).

Notably, the neural processing of RSE in the previous study was analyzed by traditional univariate analyses, which seek every single voxel that showed a significant difference in activation strength across different psychological states. However, 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). Multivariable pattern analysis (MVPA), mainly referred to as multivariate pattern classification, is a subset of machine learning statistical methodology that is known to be sensitive in detecting different psychological, cognitive or perceptual status (Sapountzis et al., 2010; Jimura and Poldrack, 2012; Wagner et al., 2019). Some studies have already leveraged MVPA methods to differentiate the unique distributed neural patterns of guilt, autobiographical memory, and somatic and vicarious pain (Rissman et al., 2016; Krishnan et al., 2016; Yu et al., 2020). Also, MVPA has demonstrated an advantage in identifying neural evidence for the validity of the measurement of implicit self-esteem (Izuma et al., 2018).

The current study

Given that self-esteem is a higher-order personality disposition and its processing is complicated (Chavez and Heatherton, 2014), the ability to identify a stable signature of RSE could contribute to our knowledge of relational self-worth. Instead of identifying the neural patterns of RSE by distinguishing it from loosely control conditions, here, we used multivariate pattern classification to investigate the widely distributed brain signature of RSE, in comparison with the PSE to eliminate unrelated processing, such as that of social desirability and self-presentation motivation (Greenwald and Satow, 1970; Baumeister et al., 1989).

It has been revealed that RSE comprises both one's value in relationships with significant others and the value of those significant others (Du et al., 2012, 2017). Regarding one's value in relation to significant others, it is similar to contextual self-referential processing where individuals adopt a first-person perspective to evaluate their traits relative to others (Chiao et al., 2009, 2010). Neuroimaging studies strongly implicate the CMS, including the mPFC (ventral and dorsal), cingulate gyrus (anterior and posterior) and precuneus (Van et al., 2010; Qin and Northoff, 2011) in mediating one's ability to consciously reflect about ourselves, that is, the self-referential processing (Northoff et al., 2006). Unlike the value of self, evaluation of the value of significant others requires not only an introspective approach but also the involvement of a perspective-taking, similar to the other-referential processing (ORP) (Frith and Frith, 2003; Gallagher and Frith, 2003). A recent review suggested that the CMSs, especially the dorsal regions of the mPFC, the posterior cingulate gyrus and the precuneus, were recruited during the processing of ORP (e.g. mother, father and spouse) (Wagner et al., 2012, 2019; Han et al., 2016). Meanwhile, TPJ was mainly responsible for processing mental inferences about others (Frank et al., 2009; Overwalle and Baetens, 2009). Overall, as the RSE is characterized as more prominently involving cognitive processing than PSE (Rosenberg et al., 1995; Pullmann and Allik, 2008), we hypothesize that the multivariable patterns of CMS and TPJ could differentiate RSE from PSE.

To this end, two experiments were designed to determine the neural patterns of RSE relative to PSE in the current study. In study 1, we trained a support vector machine (SVM) classifier to determine the brain patterns that were elicited by the evaluation of RSE, relative to PSE. For the sake of the constraint of

feature space and avoidance of overfitting, brain patterns were obtained based on the parcellation scheme of the Brainnetome atlas (Fan *et al.* 2016), which defines 246 independent regions of interest (ROIs) across the whole brain as a regional mask in the classification model. In study 2, we conducted a validation test in another independent sample to examine the reliability of our findings, in which brain patterns identified in study 1 were used to decode the neural activities related to RSE.

Study 1

Methods

Participants. Information about the participants has been previously reported in Li *et al.* (2019). Briefly, 41 right-handed healthy university students were recruited (20 males and 21 females, 18–28 years). All participants had normal or corrected-to-normal vision, and none reported any history of psychiatric or neurological disorders. Participants provided informed consent before the formal experiments, and the experiments were approved by the Ethics Committee of the University.

Stimuli and procedure. An adapted self-referential paradigm was applied in study 1. Three types of experimental stimuli were used, including RSE, PSE and semantic. In specific, stimulus on RSE was adapted from the Relational Self-Esteem Scale (Du *et al.*, 2012), which measures one's value in relationships with significant others (e.g. 'I feel I have much to offer to my family' and 'I am a worthy member of my circle of friends'), or the value of those significant others (e.g. 'Overall, my circle of friends is considered good by others' and 'My family is proud of me'). An equal number of trials were used to measure one's value and the value of significant others. The stimulus pertaining to PSE was adapted from the Rosenberg Self-Esteem Scale (Rosenberg, 1965), which measures an individuals' self-worth in general (e.g. 'I take a positive attitude toward myself' and 'On the whole, I am satisfied with myself'). The semantic stimulus was added to serve as baseline control (e.g. 'Yangzi River is the most famous river in China'). For each stimulus, participants rated to what extent they agreed with the sentence by pressing one of four buttons that mimicked a Likert scale from 1 (strongly disagree) to 4 (strongly agree). In the current analysis, we only use the RSE and PSE to detect the neural signature of RSE, relative to PSE.

Besides, we adopted an event-related design in the present study. There were four runs during the functional Magnetic Resonance Imaging (fMRI) scanning, and each run comprised 30 trials. Every 10 trials constitute a condition in each run. To eliminate the influence of order effect, we randomized the order of runs and 30 trials within one run for each participant. On each trial, the stimulus was presented in the middle of a grey screen for 4 s, followed by a jittered fixation period of 2, 4 or 6 s. The illustration of one run is shown in Figure 1.

fMRI data acquisition. Functional and anatomical whole-brain images were acquired using a 3T Siemens TRIO MRI scanner. Functional data comprised 1296 volumes functional images acquired with T2*-weighted gradient echo planar imaging (EPI) sequence were collected from each subject. We obtained 32 echo-planar images per volume, sensitive to BOLD contrast (Repetition Time (TR) = 2000 ms, Echo Time (TE) = 30 ms, 3 mm × 3 mm × 192 mm). Slices were acquired in an interleaved

order and were oriented parallel to the AC–PC plane with a thickness of 3 mm and a gap of 0.99 mm. High-resolution T1-weighted 3D fast-field echo (FFE) sequences were obtained for anatomical reference (176 slices, TR = 1900 ms, TE = 2.52 ms, slice thickness = 1 mm, Field of View (FOV) = 256 mm × 256 mm, voxel size = 1 mm × 1 mm × 1 mm).

fMRI preprocessing and first-level analysis. The details of preprocessing have been described previously (Li *et al.*, 2019). Briefly, neuroimaging data were processed and analyzed using SPM8 software (Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK). Preprocessing of the fMRI data was conducted using the Data Processing Assistant for Resting-State fMRI (DPARSF) (Yan and Zang, 2010). The functional scans were adjusted for slice timing, realigned to the first volume, co-registered to the anatomical images, normalized to a standard Montreal Neurological Institute (MNI) template and spatially smoothed with a 6 mm Full-Width and Half-Maximum (FWHM) Gaussian kernel. First-level effects were estimated by creating a general linear model, which incorporated four conditions (RSE, PSE, semantic and fixation) convolved with the canonical hemodynamic response function and six movement parameters as covariates of noninterest. Then, a high-pass temporal filter with a cutoff period of 128 s was applied.

For each of the participants, the estimated beta images corresponding to each condition of interest (i.e. RSE and PSE) and condition of no interest (i.e. semantic and fixation) were produced across runs. Notably, we used the smoothed fMRI data for MVPA based on previous research showing that smoothing can improve decoding performance when large-scale activation patterns are assumed (Beeck, 2010; Krishnan *et al.*, 2016).

Multivariate pattern analysis. The MVPA was processed by the following steps: (i) the extraction of beta maps of self-esteem, (ii) formation of weight maps of self-esteem, (iii) formation of importance maps and (iv) prediction of RSE rating.

Extraction of beta maps of self-esteem. We used the maps of whole-brain 246 regions (246 templates) to present topographic information, which parcellates the brain into 246 ROI (Fan *et al.*, 2016). The averaged beta images corresponding to RSE and PSE across runs from each ROI for each subject were constructed, resulting in 246 averaged beta values per subject for each condition. Considering that the feature dimensions (246) of the whole brain are acceptable and to prevent the loss of information during feature selection (Tian *et al.*, 2011; Liu *et al.*, 2015), we did not perform feature selection.

Formation of weight maps of self-esteem. Beta images from the RSE and PSE for 41 participants were used to train a classifier. SVM algorithm was applied as a classifier for machine learning in kernel matrix (LIBSVM, http://www.csie.ntu.edu.tw/~cjlin/libsvm/) (slack parameter C = 1 was chosen a priori). SVM had been considered to be a frequently used pattern classifier in previous analyses, which allowed the classification of individual observations into different groups or classes based on high-dimensional data (Cortes and Vapnik, 1995; Haynes and Rees, 2006; Fonseca *et al.*, 2007) and can effectively prevent the occurrence of overfitting. In classification analysis, we adopt the leave-one-subject-out cross-validation scheme. In specific, we trained the classifier on the pooled data from all but one participant and tested the ability of this classifier to predict the condition labels of brain patterns measured in the eld-out partic-

ipant. This leave-one-participant-out cross-validation scheme was iterated until each participant's data had served as the test set.

The primary classification performance metric was the area under the curve (AUC). This measure, widely used in the machine learning literature (Vilares *et al.*, 2017; Chow *et al.*, 2018) and considered more informative than overall accuracy (Bradley, 1997), can be interpreted as the probability that a randomly chosen member of one class has a smaller estimated probability of belonging to the other class than a randomly chosen member of the other class. In other words, the AUC indexes the mean accuracy with which a randomly chosen pair of Class A and Class B trials could be assigned to their correct class (chance performance is 0.5, and perfect performance is 1.0). The receiver-operating characteristic (ROC) curves, from which the AUCs are derived, reflected classifier performance within personal and RSE networks. To obtain meaningful statistical results in classification performance, we used a non-parametric permutation test to examine the null hypothesis that there was no statistical discrimination between PSE and RSE. We created 1000 randomly shuffled permutations of labels (scramble 1000 self-esteem labels in the classifier and then randomly sort them for prediction) and ran the SVM using the permutated data in each region to obtain accuracy under the classifier. To obtain the *P*-value for a particular accuracy value, we calculated the proportion of iterations with an accuracy higher than the classifier accuracy. This null hypothesis corresponds to the *P*-value of accuracy at the level of non-significance ($P > 0.05$).

Formation of importance maps. Weight maps for all regions were calculated during a classifier training cycle, which represent the contribution of a particular region to discriminating relational and PSE and were averaged across each of the leave-one-participant-out cross-validation iterations. By convention, a positive weight value indicated that a region's activity magnitude on each trial was positively correlated with the probability of that trial being from Class A, whereas a negative weight value indicated increased probability of that trial belonging to Class B. Due to trial balancing and z scoring procedures, the mean activity level of each region for Class A trials was always the additive inverse mean activity level for Class B trials and was rescaled by a constant factor of 7 to aid in visualization. Mean activities of each ROI were averaged across 41 participants from the beta values of z scoring. These weights were then multiplied by each region's mean activity level for Class A trials. Regions with positive values for both activity and weight were given positive importance values, whereas regions with negative activity and weight were given negative importance values; regions for which the activity and weight had opposite signs were assigned importance values of zero (Johnson *et al.*, 2009;

McDuff *et al.*, 2009; Rissman *et al.*, 2016). To generate importance maps for all regions depicting which regions provided maximal diagnostic signals for discrimination, we adopted the threshold of mean + 1 × s.d., which were used in previous research to represent the neural signatures of RSE relative to those of PSE (Tian *et al.*, 2011; Liu *et al.*, 2015).

Prediction of RSE and PSE rating. Using neural signals of RSE to predict the subjective RSE and PSE ratings, we further examined the discriminative validity of the diagnostic brain patterns of RSE compared with those of PSE. Combined with the leave-one-subject-out cross-validation scheme, the support vector regression (SVR) algorithm was applied to predict RSE ratings, which was obtained during the evaluation of relational self-worth. Cumulatively, there were 41 × 10 neural signals of RSE to predict ratings of RSE and PSE.

Results

As can be seen in Figure 2, the cross-validation estimation of the generalization performance shows an area of 0.87864 under the ROC curve (AUC values close to 1 indicate 'perfect' classification and close to 0.5 suggest random classification) (Figure 2). Permutation tests then further indicated that SVM was able to significantly differentiate RSE and PSE ($P < 0.001$, non-parametric permutation test, $n = 1000$ permutations).

Importance maps for the RSE *vs* PSE classification based on the threshold of mean ± 1 × s.d. revealed that diagnostic regions, including the dorsal Medial Prefrontal Cortex (dMPFC), the superior/middle temporal gyrus, the TPJ, the precuneus and the PCC, were positively predictive of RSE, whereas other diagnostic regions, including the inferior/middle frontal gyrus, the inferior/superior temporal gyrus, the insula, the PCC, the caudate, the putamen and the thalamus, were positively predictive of PSE (Table 1 and Figure 3). Besides, based on the brain patterns of RSE relative to PSE, the RSE neural signals of the dMPFC, superior/middle temporal gyrus, TPJ, precuneus and PCC could neither predict subjective RSE nor PSE ratings significantly ($r = -0.0143$, $P = 0.9292$; $r = -0.103$, $P = 0.5218$).

Discussion

The results were consistent with our hypotheses and revealed that the trained classifier could distinguish RSE and PSE with good performance, implying the evaluation of RSE and PSE involves distinct processing. Specifically, the diagnostic neural patterns of RSE relative to those of PSE were widely distributed in the superior/middle temporal gyrus, the precuneus, the PCC, the dmPFC and the TPJ. In particular, social cognitive networks such as the TPJ and the dmPFC were responsible for the evaluation
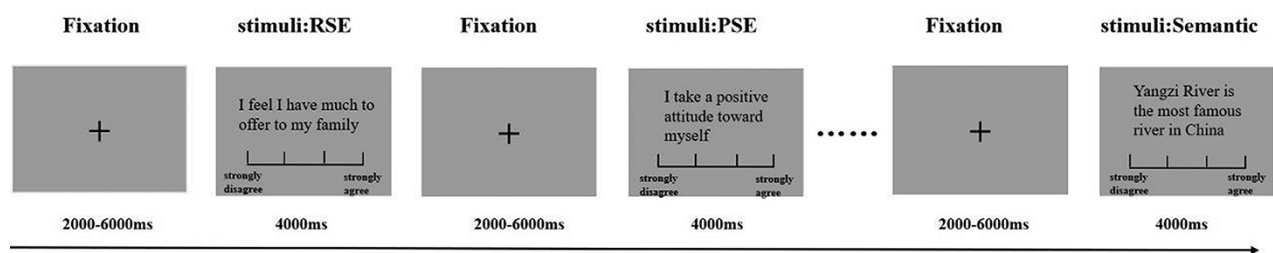

Fig. 1. Illustration of one run. There were 30 trials in each run. On each trial, a stimulus was presented in the middle of the screen for 4 s and was followed by a random jittered fixation period of 2, 4 or 6 s.

Table 1. Brain regions that positively predict RSE and PSE

Classification	Anatomical region	L/R	x	y	z
Positively predict RSE					
	Superior frontal gyrus	L	−11	49	40
	dMPFC	R	8	58	13
	Superior temporal gyrus	L	−55	−3	−10
	Middle temporal gyrus	L	−53	2	−30
	Middle temporal gyrus	L	−58	−20	−9
	TPJ	L	−47	−65	26
		R	53	−54	25
	Precuneus/PCC	L	−6	−55	34
	Precuneus/PCC	R	6	−54	35
	PCC	L	−8	−47	10
Positively predict PSE					
	Inferior/middle frontal gyrus	L	−41	41	16
		R	42	42	14
	Precentral gyrus	L	−26	−25	63
	Superior temporal gyrus	L	−54	−32	12
	Inferior temporal gyrus	R	61	−40	−17
	Inferior parietal gyrus	R	57	−44	38
	Insula	L	−38	−4	−9
		R	39	−2	−9
		L	−38	5	5
	PCC	L	−7	−23	41
		R	6	−20	40
	Lateral occipital cortex	L	−18	−99	2
	Striatum (ventral caudate)	L	−12	14	0
	Striatum (dorsal caudate)	L	−14	2	6
	Dorsolateral putamen	L	−28	−5	2
	Thalamus	L	−18	−13	3

L and R refer to left and right hemispheres; x, y and z refer to MNI coordinates that come from the Brainnetome atlas.
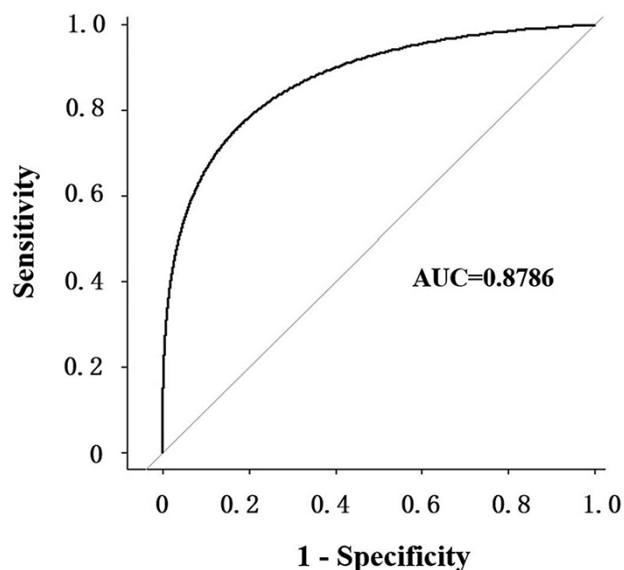

Fig. 2. Classifier performance from across-participant analysis distinguishes relational and personal self-esteem using a leave-one-participant-out scheme. The classifier performance is quantified using ROC curve analysis (AUC = 0.87864). The gray line represents chance-level classification (AUC = 0.5).

of the value of significant others. These results informed that compared with PSE, more cognitive components were prominent during the evaluation of RSE. To examine the reliability of the diagnostic neural patterns of RSE, we further conducted a validation test in another independent sample; the diagnostic neural patterns of RSE identified in study 1 were used to decode the relational self-worth in study 2.

Study 2

Methods

Participants. There were 50 right-handed healthy participants in the formal analysis (27 females, mean ages ± s.d. = 20.16 ± 1.50 years). Written informed consent was obtained from all participants following procedures approved by the Ethics Committee of the University. No participants reported a history of psychiatric or neurological illnesses, head injury or alcohol/drug use. Participants were compensated for their attendance.

Stimuli and procedure. The experimental stimulus of PSE and RSE in study 2 were the same as study 1, which were adapted from the Relational Self-Esteem Scale (Du *et al.*, 2012) and Rosenberg Self-Esteem Scale (Rosenberg, 1965). Besides, the remaining materials were associated with collective self-esteem (Luhtanen and Crocker, 1992) and semantics ('Reading is a good habit'). In
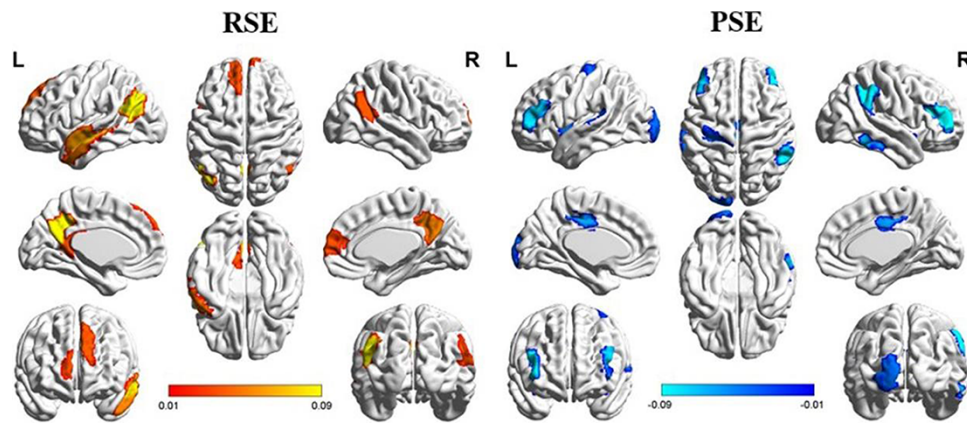
Fig. 3. Classification importance maps. Group-averaged maps of classifier importance values are shown for binary classification. Specifically, the importance values of brain regions were higher than means ± 1s.d. of classification importance values. Warm colors indicate regions for which increased activity biased the classifier to predict PSE, and cool colors indicate the regions for which increased activity biased the classifier to predict RSE. Importance values are shown in raw arbitrary units.

the current analysis, we used neural activities of RSE and PSE to assess whether the neural signature decodes the RSE activities. Participants need to determine whether they agreed with these sentences on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree).

fMRI data acquisition. The fMRI data acquisition parameters were the same as study 1. Images were acquired using a 3T Siemens TRIO MRI scanner. Functional data comprised 1296 volumes acquired with a T2*-weighted gradient EPI sequence. A total of 32 echo-planar images per volume, sensitive to BOLD contrast (TR = 2000 ms, TE = 30 ms, 3 mm × 3 mm × 192 mm), were obtained. Slices were acquired in an interleaved order and oriented parallel to the AC–PC plane, with a 0.99 mm gap. High-resolution T1-weighted 3D FFE sequences were obtained for anatomical reference (176 slices, TR = 1900 ms; TE = 2.52 ms; slice thickness = 1 mm; FOV = 256 mm 256 mm; voxel size = 1 mm × 1 mm × 1 mm).

fMRI preprocessing and first-level analysis. Preprocessing of fMRI data was performed using DPARSF (Yan and Zang, 2010), and other analyses were conducted using SPM8 (Wellcome Trust Centre for Neuroimaging, University College London, UK). The following preprocessing procedures and first-level analysis were the same as those in study 1.

Validation analysis. To examine the reliability of our findings, we conducted a validation test in an independent sample. Specifically, a widely distributed neural signal for diagnosing RSE relative to PSE should discriminate against the RSE *vs* PSE in study 1 and decode the neural activities related to RSE in study 2. We extracted beta values of the brain patterns corresponding to RSE and PSE in 50 participants. In total, there were 2 × 10 averaged beta values per subject. Then, both beta values and their labels of study 2 were put into the classifier obtained in study 1. We tested the classification accuracy when the classifier distinguished RSE *vs* PSE. Finally, the predictive classification performance of the classifier was expressed by an AUC metric. We also used a non-parametric permutation test to examine whether the classifier could distinguish RSE *vs* PSE significantly. The null hypothesis corresponds to the P-value of accuracy at the level of non-significance ($P > 0.05$). We created

1000 randomly shuffled permutations of labels (scramble 1000 self-esteem labels in the classifier and then randomly sort them for prediction) and ran the SVM using the permutated data in each region to obtain accuracy under the classifier. Finally, to obtain the P-value for a particular accuracy value, we calculated the proportion of iterations with an accuracy higher than the classifier accuracy.

Prediction of RSE and PSE rating. We also examined the discriminative validity of the diagnostic brain patterns of RSE compared with those of PSE in study 2. The SVR algorithm was combined with the leave-one-subject-out cross-validation scheme to predict RSE and PSE ratings. In all, there were 50 × 10 RSE neural signals to predict ratings of RSE and PSE.

Results

Using the brain-imaging data (beta values of brain patterns), we found that the trained classifier was sensitive to discriminate RSE *vs* PSE in an independent sample, with relatively high accuracy (AUC = 0.7320, see in Figure 4). Permutations tests then further indicated that SVM could significantly differentiate RSE and PSE ($P = 0.005$, non-parametric permutation test using $n = 1000$ permutations). In addition, based on the brain patterns of RSE relative to PSE, the neural signals could neither predict subjective RSE nor PSE ratings significantly ($r = -0.0179$, $P = 0.9017$; $r = -0.0115$, $P = 0.9371$).

Discussion

Results showed that the diagnostic neural patterns of RSE relative to PSE in study 1 could sensitively detect the neural activities related to RSE in study 2. The result indicated a reliability of the brain patterns and confirmed the more prominent cognitive components during the evaluation of RSE compared with that of PSE.

General discussion

Given that self-esteem is a higher-order personality disposition with complicated processing (Chavez and Heatherton, 2014), identifying a reliable signature of RSE would contribute to our knowledge of relational self-worth (Du *et al.*, 2012; Li *et al.*,
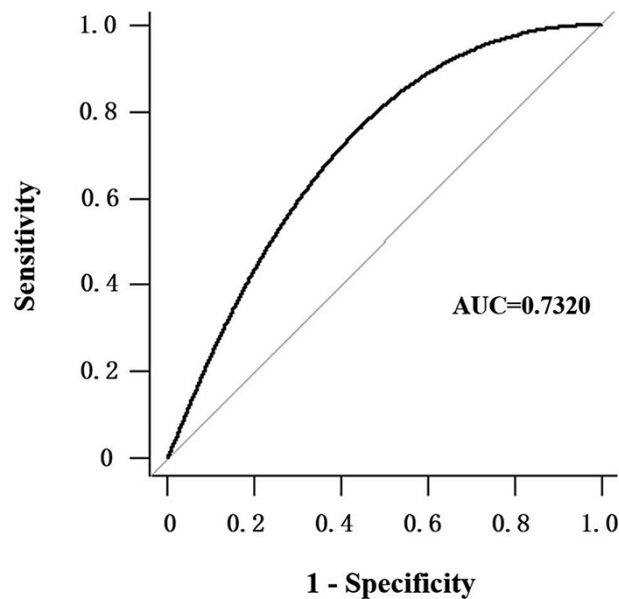
Fig. 4. Predictive classification from the validation analysis distinguishes relational self-esteem and personal self-esteem. The classifier performance was quantified using ROC curve analysis. The ROC curves differentiating RSE from PSE had an AUC of 0.7320. The gray line represents chance-level classification (AUC = 0.5).

2019). To exclude irrelevant processing, the current work utilized multivariate pattern classification to classify RSE and PSE and obtained a diagnostic signature of RSE compared with that of PSE. The results of study 1 suggested that the diagnostic signatures of RSE and PSE were involved in mentalizing and reward processing, respectively. In specific, the diagnostic signature of RSE compared with that of PSE was widely distributed in dmPFC, TPJ, superior/middle temporal gyrus, precuneus and PCC, whereas the diagnostic signature of PSE compared with that of RSE was widely distributed in the caudate, putamen, thalamus, PCC, inferior/middle frontal gyrus, inferior/superior temporal gyrus and insula. Moreover, this diagnostic signature of RSE could well distinguish RSE from PSE in study 2 and exhibited its reliability in an independent sample.

Although these peak regions parallel the previous univariate analyses (Li *et al.*, 2019), they nevertheless contribute to the understanding of the neural processing mechanisms that generate relational self-worth. Compared with PSE, the multivariate neural patterns of social cognitive networks were involved in the diagnosis of RSE. Consistent with the assumptions, the results showed that cognitive processing was prominent during the evaluation of relational self-worth. On one hand, the RSE is self-knowledge that is linked by memory to knowledge about significant others (Chen *et al.*, 2006). Activation of the specific relational situations was related to the retrieval of those memories. Brain regions like the PCC/precuneus and the middle temporal cortex play a key role in the retrieval of information from episodic and autobiographical memory (Feng *et al.*, 2018). The multivariable patterns of activity in the precuneus/PCC and middle temporal cortex were responsible for the retrieval of the relational episodic memory, which means activating the relational situations. Meanwhile, it is necessary to conceptualize beliefs and attitudes of others in the evaluation of the value of those significant others. The TPJ and the dmPFC, as key regions of the theory of mind network, play important roles in inferring

the goals, desires and beliefs of others in various social contexts (Lou *et al.*, 2004; Van, 2009; Sajonz *et al.*, 2010; Denny *et al.*, 2012). In addition, these would be involved in value assessment during the evaluation. Findings from multiple lines of inquiry revealed a role of the dmPFC in encoding the subjective value of the others (Sul *et al.*, 2015; Piva *et al.*, 2019). It may be inferred that the dmPFC is not only responsible for considering another perspective but also for calculating the value of significant others. Overall, upon analyzing the fundamental functions of the diagnostic signature of RSE, we inferred that the evaluation of relational self-worth involves the retrieval of relational episodic memory, perspective-taking and value calculation.

By training a classifier that distinguished RSE and PSE, we found that the multivariate patterns of activity in the caudate, putamen, thalamus, PCC, inferior/middle frontal gyrus, inferior/superior temporal gyrus and insula were associated with representations of the finer-grained structures of PSE relative to RSE. The caudate and the putamen, which form part of the mesolimbic dopaminergic pathway, are associated with hedonic motivation and reward (O'Doherty *et al.*, 2004; Balleine *et al.*, 2007; Peng *et al.*, 2019). Many researchers have also found these regions involve the evaluation of PSE (Chavez and Heatherton, 2014). This implies that affective components were prominent when evaluating personal self-worth, which leads to generating a reward-like emotion. Moreover, the evaluation of PSE was related to the retrieval of personal episodic memory and value calculation. The insula is related to maintaining unique self-awareness (Seth *et al.*, 2012). The multivariate neural pattern of the PCC is responsible for retrieval of the personal episodic memory (Cavanna and Trimble, 2006). In addition, inferior/middle frontal gyrus has been associated with higher-order semantic representations, such as abstract concepts and semantic contents (Cappa, 2008; Shallice and Cooper, 2013; Borghi *et al.*, 2017). This might be due to the evaluation of personal attributes and traits leading to the recruitment of these regions to retrieve more abstract semantic representation.

Taken together, by analyzing multivariable neural activities to classify RSE and PSE, our results provide neuroimaging evidence supporting the cognitively prominent processing of RSE and the affectively prominent processing of PSE. The processing of RSE comprises characteristics that define roles within the relationship and seems more concrete than PSE (Sedikides *et al.*, 2011). In this case, people would be involved in every specific context in RSE processing, which would be cognitively demanding. Conversely, regarding PSE, it comprises stable characteristics, such as traits and behaviors, which could be more abstract than RSE. In this case, people would be deeply involved in reflected emotion in the processing of PSE, which would be affectively demanding.

There are still some regions between RSE and PSE that need further explanation. Specifically, the PCC in PSE was located close to the middle portion, and the PCC/precuneus in RSE was closely located on the posterior side. The methods used in this study to explore the distinction between cognitive processing of PSE and RSE were not suitable to identify shared cognitive processing. A future study could adopt representational similarity analysis (Kriegeskorte *et al.*, 2006; Nili *et al.*, 2014) of BOLD responses between RSE and PSE, to paint a comprehensive understanding of the similarity of episodic memory between PSE and RSE. Furthermore, existing studies have found that the ventral region of the mPFC is responsible for calculating subjective values (Garvert *et al.*, 2015), especially for the self (Yankouskaya *et al.*, 2017). For

example, some researches indicate delineated functions of the ventral medial Prefrontal Cortex (vmPFC) in value-based decision-making in self-referencing decisions (McClure *et al.*, 2004; Kable and Glimcher, 2007; Levy *et al.*, 2010). However, due to the similarity of self-value computation in both RSE and PSE, we were unable to distinguish the role of the vmPFC in our classification.

This study had several limitations. First, although we have elaborated on the role of several specific brain regions for the classification of RSE and PSE, it is still important to keep in mind that the conclusions drawn from the application of the SVM pattern recognition algorithm indicate the activated brain regions should be considered as parts of an integrated regional pattern, instead of isolated regions. Thus, one must carefully interpret the contribution of individual regions in the context of the whole neural pattern. Second, the diagnostic neural patterns of RSE in our analysis were not predictive of relational self-worth ratings in study 1. One explanation is that the predictive performance of RSE may have been influenced by the sample size, which was underpowered to detect small inter-person correlations. A larger and wider sample population may show a positive predictive performance in future research. Third, caution should be taken when generalizing these neural patterns of RSE to a person with an individualistic culture. Culture likely affects the way people define themselves (Markus and Kitayama, 1991), so we speculate that people in different cultures may show distinct cognitive processing when evaluating RSE. The diagnostic signatures of RSE obtained in a more collective culture may not represent those of RSE in an individualistic culture. Future investigations may consider identifying the diagnostic neural patterns of RSE in individualistic cultures to test the cross-cultural performance of our results.

Conclusion

Using multivariate pattern classification to differentiate RSE from PSE, we obtained a stable diagnostic neural signature of RSE relative to PSE. The diagnostic signature mainly involved social cognitive networks, including the superior/middle temporal gyrus, the precuneus, the PCC, the dmPFC and the TPJ, implying that compared with PSE, cognitive components were prominent during the evaluation of RSE. Furthermore, a validation test confirmed the reliability of diagnosing brain patterns. In conclusion, by providing a reliable multivariate brain pattern for RSE relative to PSE, our findings enrich our knowledge of how RSE is generated in the brain and further supporting the cognitively prominent processing of RSE relative to PSE.

Funding

Conflict of interest

None declared.

Authors' contributions

Writing—original design: Juan Yang and Jiwen Li. Writing—data collection: Juan Yang, Mei Zeng and Mingyan Liu. Writing—data analysis: Jiwen Li, Juan Yang, Mei Zeng, Rong Li and Huafu Chen. Writing—code editing: Jiwen Li, Juan Yang, Chong Wang, Chijun Deng and Mei Zeng. Writing—original draft: Juan Yang, Jiwen Li, Xiaolin Zhao and Weiyu Hu. Writing—review and editing: Juan Yang and Jiwen Li.

References

Balleine, B.W., Delgado, M.R., Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, **27**(31), 8161–5.

Baumeister, R.F., Tice, D.M., Hutton, D.G. (1989). Self-presentational motivations and personality differences in self-esteem. *Journal of Personality*, **57**(3), 547–79.

Baumeister, R.F., Campbell, J.D., Krueger, J.I., Vohs, K.D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, **4**(1), 1–44.

Beeck, H.P.O.D. (2010). Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, **49**(3), 1943–8.

Borghi, A.M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, **143**(3), 263.

Bradley, A.P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, **30**, 1145–59.

Brown, J.D., Dutton, K.A., Cook, K.E. (2001). From the top down: self-esteem and self-evaluation. *Cognition & Emotion*, **15**(5), 615–31.

Cappa, S.F. (2008). Imaging studies of semantic memory. *Current Opinion in Neurology*, **21**(6), 669–75.

Cavanna, A.E., Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, **129**(3), 564–83.

Chavez, R.S., Heatherton, T.F. (2015). Multimodal frontostriatal connectivity underlies individual differences in self-esteem. *Social Cognitive and Affective Neuroscience*, **10**(3), 364–70.

Chen, S., Boucher, H.C., Tapias, M.P. (2006). The relational self revealed: integrative conceptualization and implications for interpersonal life. *Psychological Bulletin*, **132**(2), 151.

Chen, S., Boucher, H., Kraus, M.W. (2011). The relational self. In: *Handbook of Identity Theory & Research*, New York: Springer, 149–75.

Chiao, J.Y., Harada, T., Komeda, H., *et al.* (2009). Neural basis of individualistic and collectivistic views of self. *Human Brain Mapping*, **30**(9), 2813–20.

Chiao, J.Y., Harada, T., Komeda, H., *et al.* (2010). Dynamic cultural influences on neural representations of the self. *Journal of Cognitive Neuroscience*, **22**(1), 1–11.

Chow, T.E., Westphal, A.J., Rissman, J. (2018). Multi-voxel pattern classification differentiates personally experienced event memories from secondhand event knowledge. *NeuroImage*, **176**, 110–23.

Cortes, C., Vapnik, V. (1995). Support-vector networks. *Machine Learning*, **20**(3), 273–97.

Crocker, J., Wolfe, C.T. (2001). Contingencies of self-worth. *Psychological Review*, **108**(3), 593.

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. *Journal of Cognitive Neuroscience*, **24**(8), 1742–52.

Du, H., King, R.B., Chi, P. (2012). The development and validation of the Relational Self-Esteem Scale. *Scandinavian Journal of Psychology*, **53**(3), 258–64.

Du, H., Jonas, E., Klackl, J., Agroskin, D., Hui, E.K.P., Ma, L. (2013). Cultural influences on terror management: independent and interdependent self-esteem as anxiety buffers. *Journal of Experimental Social Psychology*, **49**(6), 1002–11.

Du, H., King, R.B., Chi, P. (2017). Self-esteem and subjective well-being revisited: the roles of personal, relational, and collective self-esteem. *PLoS One*, **12**(8), e0183958.

Eisenberger, N.I., Inagaki, T.K., Muscatell, K.A., Byrne Haltom, K.E., Leary, M.R. (2011). The neural sociometer: brain mechanisms underlying state self-esteem. *Journal of Cognitive Neuroscience*, **23**(11), 3448–55.

Fan, L., Li, H., Zhuo, J., *et al.* (2016). The human Brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral Cortex*, **26**(8), 3508–26.

Feng, C., Yan, X., Huang, W., Han, S., Ma, Y. (2018). Neural representations of the multidimensional self in the cortical midline structures. *NeuroImage*, **183**, 291–9.

Fonseca, E.S., Guido, R.C., Scalassara, P.R., Maciel, C.D., Pereira, J.C. (2007). Wavelet time-frequency analysis and least squares support vector machines for the identification of voice disorders. *Computers in Biology and Medicine*, **37**(4), 571–8.

Frewen, P.A., Lundberg, E., Jean, Brimson-Théberge, M., & Théberge, J. (2013). Neuroimaging self-esteem: a fMRI study of individual differences in women. *Social Cognitive and Affective Neuroscience*, **8**(5), 546–55.

Frith, U., Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **358**(1431), 459–73.

Gallagher, H.L., Frith, C.D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, **7**(2), 77–83.

Garvert, M.M., Moutoussis, M., Kurth-Nelson, Z., Behrens, T.E., Dolan, R.J. (2015). Learning-induced plasticity in medial prefrontal cortex predicts preference malleability. *Neuron*, **85**(2), 418–28.

Georgopoulos, A., Schwartz, A., Kettner, R. (1986). Neuronal population coding of movement direction. *Science*, **233**(4771), 1416–9.

Greenwald, H.J., Satow, Y. (1970). A short social desirability scale. *Psychological Reports*, **27**(1), 131–5.

Han, S., Ma, Y., Wang, G. (2016). Shared neural representations of self and conjugal family members in Chinese brain. *Culture and Brain*, **4**(1), 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**(5539), 2425–30.

Haynes, J.D., Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, **7**(7), 523.

Izuma, K., Kennedy, K., Fitzjohn, A., Sedikides, C., Shibata, K. (2018). Neural activity in the reward-related brain regions predicts implicit self-esteem: a novel validity test of psychological

measures using neuroimaging. *Journal of Personality and Social Psychology*, **114**(3), 343–57.

Jimura, K., Poldrack, R.A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, **50**(4), 544–52.

Johnson, J.D., McDuff, S.G., Rugg, M.D., Norman, K.A. (2009). Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron*, **63**(5), 697–708.

Kable, J.W., Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, **10**(12), 1625–33.

Kosslyn, S.M. (1999). If neuroimaging is the answer, what is the question? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **354**(1387), 1283–94.

Kriegeskorte, N., Goebel, R., Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(10), 3863–8.

Krishnan, A., Woo, C.W., Chang, L.J., *et al.* (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *Elife*, **5**, e15166.

Leary, M.R., Baumeister, R.F. (2000). The nature and function of self-esteem: sociometer theory. *Advances in Experimental Social Psychology*, **32**, 1–62.

Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, **103**(2), 1036–47.

Li, J., Liu, M., Peng, M., Jiang, K., Chen, H., Yang, J. (2019). Positive representation of relational self-esteem versus personal self-esteem in Chinese with interdependent self-construal. *Neuropsychologia*, **134**, 107195.

Liu, F., Xie, B., Wang, Y., *et al.* (2015). Characterization of post-traumatic stress disorder using resting-state fMRI with a multi-level parametric classification approach. *Brain Topography*, **28**(2), 221–37.

Lou, H.C., Luber, B., Crupain, M., *et al.* (2004). Parietal cortex and representation of the mental self. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(17), 6827–32.

Luhtanen, R., Crocker, J. (1992). A collective self-esteem scale: self-evaluation of one's social identity. *Personality & Social Psychology Bulletin*, **18**(3), 302–18.

Markus, H.R., Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review*, **98**(2), 224.

McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, **306**(5695), 503–7.

McDuff, S.G., Frankel, H.C., Norman, K.A. (2009). Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. *Journal of Neuroscience*, **29**(2), 508–16.

Moran, J., Macrae, C., Heatherton, T.F., Wyland, C., Kelley, W.M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, **18**(9), 1586–94.

Murray, R.J., Debbané, M., Fox, P.T., Bzdok, D., Eickhoff, S.B. (2015). Functional connectivity mapping of regions associated with self-and other-processing. *Human Brain Mapping*, **36**(4), 1304–24.

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology*, **10**(4), e1003553.

Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *NeuroImage*, **31**(1), 440–57.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**(5669), 452–4.

Ochsner, K.N., Lieberman, M.D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, **56**(9), 717–34.

Overwalle, F.V. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, **30**(3), 829–58.

Overwalle, F.V., Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, **48**(3), 564–84.

Pan, W., Liu, C., Qian, Y., Yan, G., Yin, S., Chen, A. (2016). The neural basis of trait self-esteem revealed by the amplitude of low-frequency fluctuations and resting state functional connectivity. *Social Cognitive and Affective Neuroscience*, **11**(3), 367–76.

Peng, M., Wu, S., Shi, Z., *et al*. (2019). Brain regions in response to character feedback associated with the state self-esteem. *Biological Psychology*, **148**, 107734.

Piva, M., Velnoskey, K., Jia, R., Nair, A., Levy, I., Chang, S.W. (2019). The dorsomedial prefrontal cortex computes task-invariant relative subjective value for self and other. *Elife*, **8**, e44939.

Pullmann, H., Allik, J. (2008). Relations of academic and general self-esteem to school achievement. *Personality and Individual Differences*, **45**(6), 559–64.

Qin, P., Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage*, **57**(3), 1221–33.

Rissman, J., Chow, T.E., Reggente, N., Wagner, A.D. (2016). Decoding fMRI signatures of real-world autobiographical memory retrieval. *Journal of Cognitive Neuroscience*, **28**(4), 604–20.

Rosenberg, M. (1965). *Society and the Adolescent Self-image*, Princeton, NJ: Princeton University Press.

Rosenberg, M., Schooler, C., Schoenbach, C., Rosenberg, F. (1995). Global self-esteem and specific self-esteem: different concepts, different outcomes. *American Sociological Review*, **60**(1), 141–56.

Sajonz, B., Kahnt, T., Margulies, D.S., *et al*. (2010). Delineating self-referential processing from episodic memory retrieval: common and dissociable networks. *NeuroImage*, **50**(4), 1606–17.

Sapountzis, P., Schluppeck, D., Bowtell, R., Peirce, J.W. (2010). A comparison of fMRI adaptation and multivariate pattern classification analysis in visual cortex. *NeuroImage*, **49**(2), 1632–40.

Sedikides, C., Gaertner, L., O'Mara, E.M. (2011). Individual self, relational self, collective self: hierarchical ordering of the tripartite self. *Psychological Studies*, **56**(1), 98–107.

Seth, A.K., Suzuki, K., Critchley, H.D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, **2**, 395.

Shallice, T., Cooper, R.P. (2013). Is there a semantic system for abstract words? *Frontiers in Human Neuroscience*, **7**, 175.

Sul, S., Tobler, P.N., Hein, G., *et al*. (2015). Spatial gradient in value representation along the medial prefrontal cortex reflects individual differences in prosociality. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(25), 7851–6.

Tajfel, H., Turner, J.C. (1986). The social identity theory of intergroup behavior. In: Worchel, S., Austin, W.G., editors. *Psychology of Intergroup Relations*, 2nd edn, Chicago, IL: Nelson-Hall, 7–24.

Tian, L., Wang, J., Yan, C., He, Y. (2011). Hemisphere- and gender-related differences in small-world brain networks: a resting-state functional MRI study. *NeuroImage*, **54**(1), 191–202.

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. *Neuroscience and Biobehavioral Reviews*, **34**(6), 935–46.

van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, **30**(3), 829–58.

Vilares, I., Wesley, M.J., Ahn, W.Y., *et al*. (2017). Predicting the knowledge–recklessness distinction in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(12), 3222–7.

Wagner, D.D., Haxby, J.V., Heatherton, T.F. (2012). The representation of self and person knowledge in the medial prefrontal cortex. *Wiley Interdisciplinary Reviews Cognitive Science*, **3**(4), 451–70.

Wagner, D.D., Chavez, R.S., Broom, T.W. (2019). Decoding the neural representation of self and person knowledge with multivariate pattern analysis and data-driven approaches. *Wiley Interdisciplinary Reviews Cognitive Science*, **10**(1), e1482.

Wagner, J. (2009). How relational is self-evaluation? Self-esteem and social relationships across life span and family situations. Doctoral dissertation, Friedrich-Alexander-University of Erlangen-NÜrnberg.

Yan, C., Zang, Y. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience*, **4**, 13.

Yang, J., Dedovic, K., Chen, W., Zhang, Q. (2012). Self-esteem modulates dorsal anterior cingulate cortical response in self-referential processing. *Neuropsychologia*, **50**(7), 1267–70.

Yang, J., Xu, X., Chen, Y., Shi, Z., Han, S. (2016). Trait self-esteem and neural activities related to self-evaluation and social feedback. *Scientific Reports*, **6**(1), 20274.

Yankouskaya, A., Humphreys, G., Stolte, M., Stokes, M., Moradi, Z., Sui, J. (2017). An anterior–posterior axis within the ventromedial prefrontal cortex separates self and reward. *Social Cognitive and Affective Neuroscience*, **12**(12), 1859–68.

Yu, H., Koban, L., Chang, L.J., *et al*. (2020). A generalizable multivariate brain pattern for interpersonal guilt. *Cerebral Cortex*, **30**(6), 3558–72.