OXFORD

Social Cognitive and Affective Neuroscience, 2023, **18**(1), 1–10

DOI: https://doi.org/10.1093/scan/nsad002
Advance Access Publication Date: 25 January 2023
Original Manuscript

Connectome-based predictive modeling of trait forgiveness

Jingyu Li,[1,2] Jiang Qiu,[3,4] and Haijiang Li[1,2]

[1]Department of Psychology, Shanghai Normal University, Shanghai 200234, China
[2]The Research Base of Online Education for Shanghai Middle and Primary Schools, Shanghai 200234, China
[3]Key Laboratory of Cognition and Personality (SWU), Ministry of Education, Chongqing 400715, China
[4]Faculty of Psychology, Southwest University, Chongqing 400715, China
Correspondence should be addressed to Haijiang Li, Department of Psychology, Shanghai Normal University, No. 100 Guilin Rd., Xuhui District, Shanghai 200234, China. E-mail: haijiangli@shnu.edu.cn

Abstract

Forgiveness is a positive, prosocial manner of reacting to transgressions and is strongly associated with mental health and well-being. Despite recent studies exploring the neural mechanisms underlying forgiveness, a model capable of predicting trait forgiveness at the individual level has not been developed. Herein, we applied a machine-learning approach, connectome-based predictive modeling (CPM), with whole-brain resting-state functional connectivity (rsFC) to predict individual differences in trait forgiveness in a training set (dataset 1, $N = 100$, 35 men, 17–24 years). As a result, CPM successfully predicted individual trait forgiveness based on whole-brain rsFC, especially via the functional connectivity of the limbic, prefrontal and temporal areas, which are key contributors to the prediction model comprising regions previously implicated in forgiveness. These regions include the retrosplenial cortex, temporal pole, dorsolateral prefrontal cortex (PFC), dorsal anterior cingulate cortex, precuneus and dorsal posterior cingulate cortex. Importantly, this predictive model could be successfully generalized to an independent sample (dataset 2, $N = 71$, 17 men, 16–25 years). These findings highlight the important roles of the limbic system, PFC and temporal region in trait forgiveness prediction and represent the initial steps toward establishing an individualized prediction model of forgiveness.

Keywords: forgiveness; connectome-based predictive modeling; resting-state functional connectivity

Introduction

Forgiveness is a positive, prosocial manner of reacting to harm and reducing stress brought about by unforgiveness (i.e. anger, hostility, etc.); its process involves prosocial changes in emotion, cognition and behavior (Toussaint and Friedman, 2008). Trait forgiveness is the disposition to forgive interpersonal transgressions over time and across situations (Berry *et al.*, 2005). Higher trait forgiveness levels indicate a greater tendency to respond to transgressions in a prosocial manner. Previous studies found that forgiveness had a significant negative association with neuroticism (Brose *et al.*, 2005), depression (Gençoğlu *et al.*, 2018) and rumination (Berry *et al.*, 2005). Moreover, forgiveness strongly correlates with physical health; individuals with low levels of forgiveness experience higher blood pressure, heart rate, perceived stress and loneliness (Lawler-Row *et al.*, 2011). Exploring individual differences and their neural bases may help assess individuals confronting continuous negative emotions, further improving their mental health (Griffin *et al.*, 2015).

Prior meta-analyses have proposed that forgiveness encompasses at least three psychological macro-processes that are supported by distinct brain networks involved in cognitive control, perspective taking and social valuation (Fourie *et al.*, 2020). For example, imaging studies have revealed that reduced functional connectivity between the medial prefrontal cortex (mPFC) and dorsal anterior cingulate cortex (dACC)—key nodes of cognitive control—is associated with increased acceptance of unfair offers from transgressors, which indicates forgiveness (Fatfouta *et al.*, 2016). Furthermore, the dorsolateral prefrontal cortex (dlPFC) plays an important role in inhibiting unwanted emotional responses, which must be suppressed to act in an unforgiving manner (Maier *et al.*, 2018, 2019). Perspective taking is crucial for victims to understand the wrongdoer's behavior and intention and thus further consider forgiving the wrongdoer, and some studies have identified activation in areas associated with perspective taking, including the temporoparietal junction, mPFC, precuneus and posterior cingulate cortex (PCC) (Farrow *et al.*, 2005; Ohtsubo *et al.*, 2018), which also strongly correlate with forgiveness. Consistently, resting-state brain activity variation in mentalizing regions has been associated with individual differences in the tendency to forgive (Li and Lu, 2017).

Received: 30 July 2022; **Revised:** 29 December 2022; **Accepted:** 24 January 2023

Ricciardi et al. (2013) found that forgiveness is associated with positive emotional states and the strength of the connection between the precuneus and inferior parietal lobule was significantly correlated with participants' subjective relief during forgiveness. Lastly, when interacting with the potential harm inflictor, people combine information about relationship values and exploitation risk with a harm inflictor to arrive at a decision (Burnette et al., 2012; McCullough et al., 2013); this relies fundamentally on the ventromedial PFC, including the anterior PFC, medial sector of the orbitofrontal cortex and subgenual ACC (Rudebeck et al., 2008).

Predicting individual mental traits and behavioral dispositions from brain imaging data through machine-learning methods [e.g. connectome-based predictive modeling (CPM)] is a rapidly evolving field in neuroscience and may provide the basis for more objective and valid assessments of personal aptitudes, attitudes and other mental characteristics (Eickhoff and Langner, 2019). CPM has been tested as a data-driven approach to account for interindividual variability in functional brain networks (Cai et al., 2020; Rutherford et al., 2020; Yang et al., 2021). It can help develop prediction models of brain–behavior relationships that can detect individual variability more accurately. With the continuing increase in research using this method, researchers are increasingly advocating larger samples, external validation and the combination of multiple cross-validation (CV) methods to improve the generalizability of the model (Yeung et al., 2022).

Therefore, the current study aimed to present initial efforts in this direction by making individualized predictions of forgiveness from intrinsic whole-brain functional connectivity. CPM with multiple CV methods was implemented and tested on an external dataset to predict individual trait forgiveness using whole-brain resting-state functional connectivity (rsFC). Rather than testing a specific hypothesis, CPM can implement more holistic measures with whole-brain analyses. Moreover, an inspection of the network neuroanatomy can aid the hypothesis generation in future studies (Ren et al., 2021). Hence, we inspected the rsFC of the connections that make up the 'forgiveness connectome' to determine a framework for hypothesis testing in future research. Based on previous findings, we expected that individual differences in trait forgiveness would be predicted by functional connectivity across distributed networks, particularly those implicated in cognitive control, perspective taking and social valuation.

Materials and methods

Participants

As CPM is particularly informative when the predictive value of the findings from a discovery dataset can be tested on a separate dataset (Ren et al., 2021), two samples of participants were recruited from Southwest University in Chongqing, China. Dataset 1 was used as the discovery dataset and comprised 121 participants. Participants with missing imaging data, incomplete psychological assessment or excessive head motion (defined as >2.5 mm translation or >2.5° rotation during the run) were excluded. Finally, 100 participants were retained as the discovery dataset (35 men; 19.91 ± 1.38 years old, range: 17–24 years). Dataset 2 was used as the external dataset and comprised 93 participants. Lastly, 71 additional participants were retained as the validation dataset for external validation (17 men; 19.86 ± 2.02 years old, range: 16–25 years).

All participants were free of neurological impairments and psychiatric disabilities. The human procedures were approved by the Southwest University Brain Imaging Center Institutional Review Board. Participants provided written informed consent before the study and were paid after the experiment.

Assessment of trait forgiveness

Trait forgiveness was assessed in both samples using the Trait Forgiveness Scale (TFS; Berry et al., 2005). It consists of 10 items, and each item is scored on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Items such as 'People close to me probably think I hold a grudge too long' and 'I can forgive a friend for almost anything' indicated the respondents' proneness to forgive interpersonal transgressions. A higher score indicated a stronger tendency to forgive others. The Cronbach alpha value of the scale of the current research was 0.798 in dataset 1 and 0.633 in dataset 2.

Functional magnetic resonance imaging image acquisition

Images were acquired using a 3 T Trio scanner (Siemens Medical Systems, Erlangen, Germany) at the Southwest University Brain Imaging Center. Resting-state scanning consisted of 242 gradient echo-planar imaging (EPI) volumes using the following parameters: repetition time = 2000 ms; echo time = 30 ms; slices = 32; thickness = 3 mm; resolution matrix = 64×64; flip angle = 90°; slice gap = 1 mm; field of view = 192×192 mm^2 and voxel size = $3.4 \times 3.4 \times 3.4$ mm^3.

All participants underwent 8-min resting-state functional magnetic resonance imaging (fMRI) scan, during which they were required to remain still, close their eyes, remain awake and think of nothing specific. No participant reported sleeping during scanning.

fMRI pre-processing

The functional images were pre-processed in Data Processing Assistant for Resting-State fMRI (DPARSF, http://rfmri.org/DPARSF; Yan and Zang, 2010) on SPM 12 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm).

The 10 initial volumes were removed to ensure steady magnetization. Then, 232 volumes remained for the slice timing to correct for intra-volume acquisition delay. The images were further realigned for head-motion correction. 14 participants were excluded from further analyses under the criteria of head motion exceeding 2.5 mm maximum translation and 2.5° rotation. The functional volumes were then normalized using EPI templates (voxel size was $3 \times 3 \times 3$ mm^3). Next, the images were spatially smoothed using a Gaussian filter to decrease spatial noise ($4 \times 4 \times 4$ mm^3 full width at half maximum). Subsequently, the linear trends of the time courses were removed, and band-pass filtering (0.01–0.1 Hz) was applied to the time series of each voxel to reduce the effects of low-frequency drifts and high-frequency physiological noise. Finally, nuisance covariate regression (24 Friston parameters, white matter and cerebrospinal fluid) was also applied to the volumes.

Functional connectivity construction

To improve the reliability and sensitivity of the network analyses, the static whole-brain rsFC was constructed using the Shen 268-node brain atlas, derived from a graph-theory-based parcellation algorithm with higher parcellation accuracy and spatial coherence (Shen et al., 2010, 2013). Consistent with previous studies (Beaty et al., 2018; Feng et al., 2019; Wang et al., 2021), this atlas divides the brain into 10 lobes containing 268 regions of interest (ROIs) to define network nodes, including the prefrontal lobe (46 nodes), motor lobe (21 nodes), insular lobe (7 nodes), parietal lobe (27 nodes) and temporal lobe (39 nodes). For each participant, node-by-node pairwise correlations (Pearson's r) were computed, and Fisher's r-to-z transformation was implemented to improve the normality of the correlation coefficients, resulting in a 268×268 symmetric connectivity matrix in which each element represents the connectivity strength (sometimes referred to as an 'edge') between two individual nodes.

Connectome-based predictive modeling

CPM was conducted to predict TFS scores using previously published MATLAB scripts, which are freely available online (https://www.nitrc.org/projects/bioimagesuite/), within the discovery dataset of 100 participants (dataset 1). CPM contains the following main steps:

(i) Edge selection: with the rsFC matrixes and behavioral data as inputs, each edge in the rsFC matrixes correlated with the behavioral data (TFS) using Pearson's correlation or partial Pearson's correlation to avoid potential confounding effects. The most significantly correlated edges under an optimal threshold (see validation analyses) were selected as the predictive edges.

(ii) Network construction: a positive network was constructed using selected edges that were positively correlated with TFS (increased connectivity was associated with behavioral data); and a negative network was constructed using selected edges that were negatively correlated (decreased connectivity was associated with behavioral data).

(iii) Single-subject summary value: a single-subject summary value was obtained by summing the values of all edges in the rsFC matrix in the positive network and the negative network separately for each participant.

(iv) Model building: fitting the linear model between TFS scores and network strengths across the training set for both the positive and negative networks to build the positive and negative models, respectively. We subsequently used the two networks in combination to obtain a combined model.

(v) Model validation: to determine whether the predicted TFS scores generated by the previous step significantly predicted the observed scores.

Further details are provided in the following sections. For a more in-depth explanation of the CPM technique, refer to Shen *et al.* (2017).

Validation analyses

Feature selection threshold

Although we reported the main results with a threshold of $P < 0.05$, we also examined the results with five other thresholds (0.01, 0.005, 0.001, 0.0005 and 0.0001). In terms of 10 simple rules for applying predictive modeling to rsFC data (Scheinost *et al.*, 2019) and to be consistent with past work employing CPM (Beaty *et al.*, 2018; Jiang *et al.*, 2018; Feng *et al.*, 2019; Ren *et al.*, 2021), leave-one-out cross-validation (LOOCV) was used first. In the LOOCV process, $N - 1$ participants were used as the training set, and the remaining one was used as the validation sample, where N is the number of participants. In each iteration, 99 participants' rsFC matrixes and TFS scores were used as inputs to perform the aforementioned steps, resulting in a set of parameters and a model to predict the TFS score of one participant. The training and validation procedures were repeated N times such that each participant was used once as the validation participant.

The Pearson correlation coefficient (r) and mean squared error (MSE) between the actual and predicted TFS scores were used to evaluate the accuracy of the prediction. In order to determine whether the obtained metrics were significantly better than expected by chance, a permutation test was applied. To generate null distributions for significance testing, we randomly shuffled the correspondence between connectivity matrixes and behavioral variables 5000 times and reran the CPM pipeline using the shuffled data. Based on the null distribution, the P value for the leave-one-out prediction was calculated as the proportion of sampled permutations that were greater than or equal to the true prediction correlation, that is, the P value = the number of permutations that generated correlation values greater than or equal to the true correlation values/5000. Statistical significance was set at $P < 0.05$.

Different CV schemes

However, the LOOCV strategy may generate biased estimates (Kohavi, 1995; Varoquaux *et al.*, 2017), and different CV schemes were used to obtain more comprehensive and stable outcomes. Therefore, the main results were further validated using different CV schemes (i.e. 5-fold, 10-fold, and 20-fold). Taking the 10-fold as an example, all participants were grouped into 10 subsets; nine subsets were used as the training sets, and the remaining subset was used as the validation set. This procedure was repeated 10 times so that each subset was used as the validation set. Because the full dataset was randomly divided into 10 subsets, the performance might depend on the data division. Therefore, 10-fold CV was repeated 100 times to obtain the average prediction performance.

Control analyses

Control analyses were conducted to avoid the potential confounding effects of head motion, age and sex. We repeated our analyses by implementing scrubbing with the criterion of a frame-wise displacement during scanning (FD) >0.2 mm (Jenkinson *et al.*, 2002; Yan *et al.*, 2013). In these analyses, edge selection was conducted using partial Pearson correlation with the mean FD of head motion, age and sex as covariables separately.

External generalizability

CV is acceptable, but it should be noted that CV in small samples may render the models too optimistic and that external validation is the best practice (Whelan and Garavan, 2014; Yeung *et al.*, 2022). To generate a final model for application to a completely independent sample (dataset 2), we calculated the brain networks of TFS and the model parameters derived from dataset 1 to predict the TFS scores in the previously unseen dataset. In dataset 1, a total of 100 iterations were run to generate predicted TFS scores for a different left-out participant in that sample. We then took the average of the model parameters (i.e. the slopes and intercepts) to build the regression models that would predict TFS scores for dataset 2. The model performance predictive power in the external dataset was assessed by correlating the model-predicted and actual TFS scores.

Results

Behavioral results

We first examined the distribution of TFS scores (M = 31.71, s.d. = 5.556), and the one-sample Kolmogorov–Smirnov (K–S) test showed that TFS scores were normally distributed (K–S = 0.068, $P = 0.200$). The participants differed widely in their TFS scores (Figure 1A). No sex differences were found for TFS [t (98) = 0.182, $P = 0.856$]. TFS was not significantly correlated with

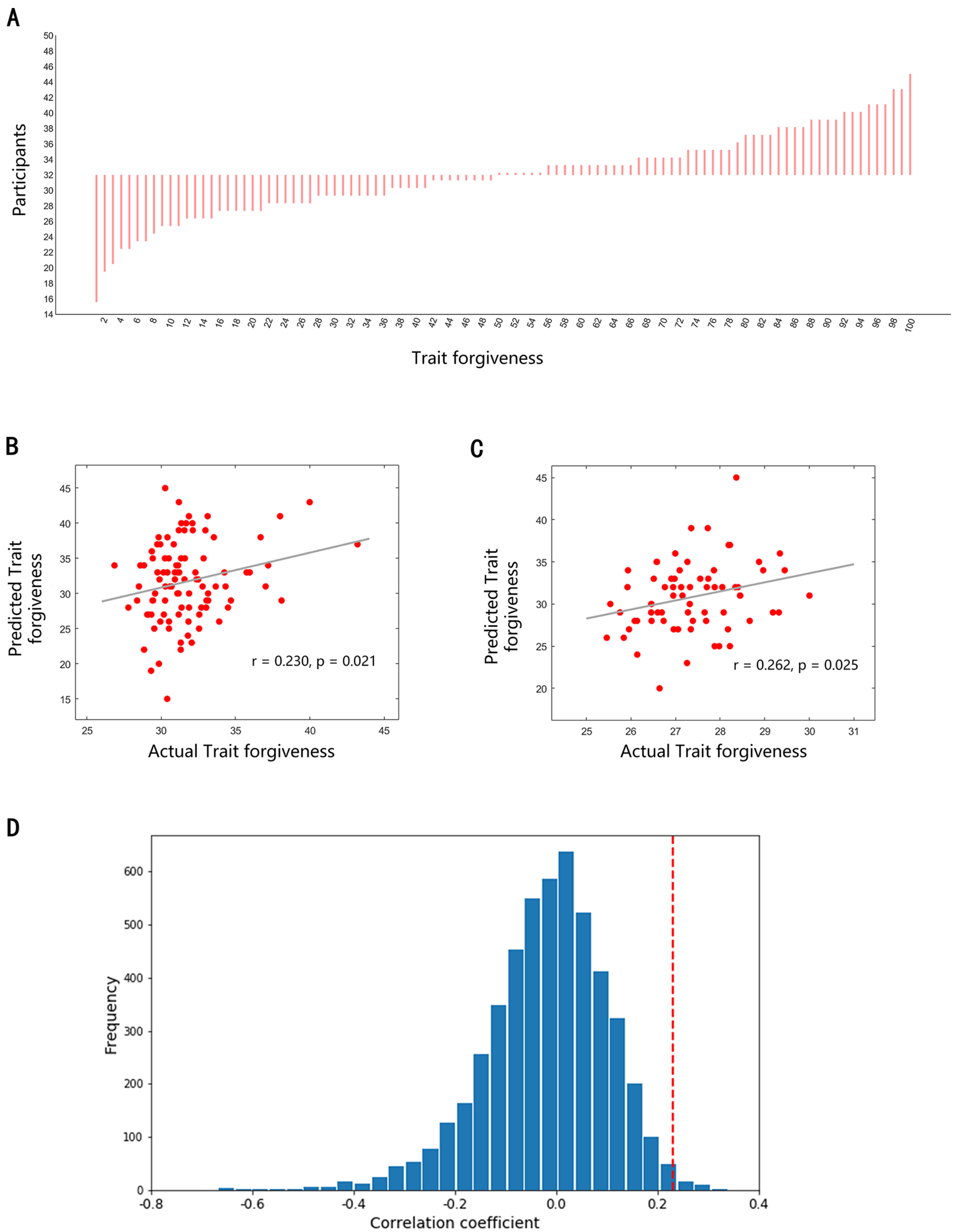

Fig. 1. Performance of the prediction model. (A) Scores of trait forgiveness across all participants. (B) Correlation between actual and predicted trait forgiveness scores using rsFC in the positive network. (C) Correlation between actual and predicted trait forgiveness scores on external dataset. (D) Permutation distribution of the correlation coefficient (r) for the prediction analysis.

participant age ($r=0.099$, $P=0.325$) or with head motion (FD) during scanning ($r=0.182$, $P=0.070$).

CPM results from dataset 1 (n = 100)

After the CPM process, we found that a positive network ($r\_pos=0.230$, $P=0.021$, MSE = 31.55; Figure 1B and D), a negative network ($r=0.260$, $P=0.009$, MSE = 29.81) and a combined network ($r=0.312$, $P=0.002$, MSE = 29.92) could predict TFS scores at a threshold of 0.05; the results at the other thresholds (0.01, 0.005, 0.001, 0.0005 and 0.0001) are displayed in Table 1.

Contributing networks in the prediction of TFS scores

As suggested by previous studies, we applied a threshold ($P<0.05$) and LOOCV to retain the most significant edges in the connectivity

Table 1. Results of different CV schemes

	Positive network			Negative network			Combined network	
	r	**P**	**MSE**	**r**	**P**	**MSE**	**r**	**P**
Threshold								
0.05	0.230	0.021	31.56	0.260	0.009	29.81	0.312	0.002
0.01	0.214	0.033	32.85	0.299	0.003	29.13	0.310	0.002
0.005	0.201	0.045	33.67	0.330	0.001	28.32	0.307	0.002
0.001	0.148	0.143	36.19	0.254	0.011	30.47	0.242	0.015
0.0005	0.141	0.162	36.19	0.246	0.014	30.86	0.233	0.020
0.0001	0.135	0.181	36.02	0.346	0.000	28.04	0.269	0.007

Abbreviations: *r*, Pearson correlation coefficient in network; *P*, probability value in network.
Note: this table shows the *r*, *P* and MSE in positive, negative and combined networks across different thresholds.

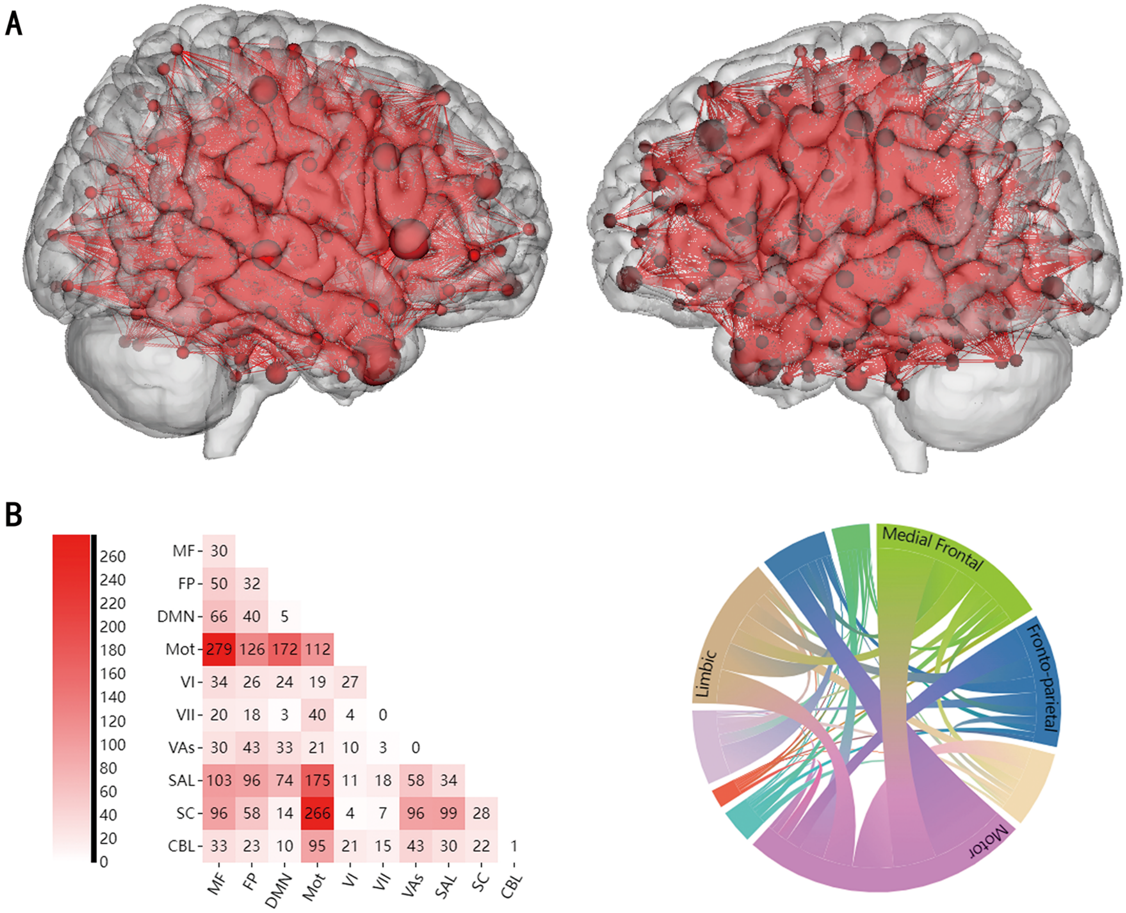

Fig. 2. Functional connections predicting trait forgiveness scores. The results were visualized using BioImage Suite (http://bioimagesuite.com/). (A) Positive networks selected by the model; increased edge weights predict higher trait forgiveness scores. (B) Connections plotted as the number of edges within and between each pair of canonical networks: in the left matrix, larger numbers in cells indicate a greater number of edges connecting nodes within and between each network; the connections are shown diagrammatically on the right.

matrixes for the following analyses. The total possible edges were defined by the atlas used in this work (Shen *et al.*, 2013), which were $(268 \times 267) \div 2 = 35778$. After LOOCV, there were 2804 edges in the positive network and 57 edges in the negative network that appeared in every iteration and were defined as the contributing network (Rosenberg *et al.*, 2016) (Figure 2A). Because the limited number of edges in the negative network could not provide stable predictions, the following analyses focused on the positive network (Feng *et al.*, 2019). Figure 2B shows the connectivity based on the number of connections within and between canonical networks for the positive network: this positive network included default mode network (DMN), frontoparietal network (FPN), motor network and limbic network, which were highly involved in the prediction.

Additionally, the 20 most highly connected nodes were located in the right temporal pole, retrosplenial cortex, (dlPFC), dACC and dorsal PCC extending to the precuneus, indicating that these nodes play a critical role in predicting trait forgiveness (Table 2, Figure 3).

Validation analyses

With different CV schemes, the performance of the predictive model was re-evaluated at the 0.05 threshold, and the results remained significant (Table 3).

Table 2. The top 20 nodes with the most connections selected by the positive prediction model

No.	Node	MNI coordinates (mm)			Lobe	Degree
51	Temporalpole (BA 38)	27.18	11.56	−39.16	R-temporal	110
227	AgrRetrolimb (BA 30)	−7.47	−42.12	13.32	L-limbic	104
263	Thalamus	−4.87	−10.34	5.83	L-subcortical	97
16	ParsOrbitalis (BA 47)	53.58	24.81	0.89	R-prefrontal	91
128	Thalamus	5.46	−9.67	5.24	R-subcortical	89
127	Thalamus	12.26	−27.74	13.5	R-subcortical	71
262	Thalamus	−9.59	−25.43	−1.42	L-subcortical	69
63	SupTempGyrus (BA 22)	61.85	−23.77	−2.81	R-temporal	67
146	dlPFC (BA 9)	−27.33	34.07	36.39	L-prefrontal	67
83	DorsalACC (BA 32)	7.84	34.68	17.09	R-limbic	64
33	PrimSensory (BA 1)	41.97	−23.38	53.41	R-motorstrip	63
225	DorsalPCC (BA 31)	−6.5	−53.94	37.44	L-limbic	63
158	PrimMotor (BA 4)	−41.59	−14.68	44.79	L-motorstrip	62
261	Putamen	−24.78	5.62	−0.08	L-subcortical	62
9	AntPFC (BA 10)	28.88	51.14	18.68	R-prefrontal	60
22	Broca-Operc (BA 44)	39.98	17.61	29.19	R-prefrontal	60
139	AntPFC (BA 10)	−18.21	56.99	−14.27	L-prefrontal	57
62	PrimAuditory (BA 41)	39.86	−25.56	14.38	R-temporal	56
155	Broca-Triang (BA 45)	−32.45	22.12	5.84	L-prefrontal	54
172	PrimSensory (BA 1)	−23.5	−31.62	63.61	L-parietal	54

Abbreviations: L, left; R, right; BA, Brodmann area; MNI, Montreal Neurological Institute.

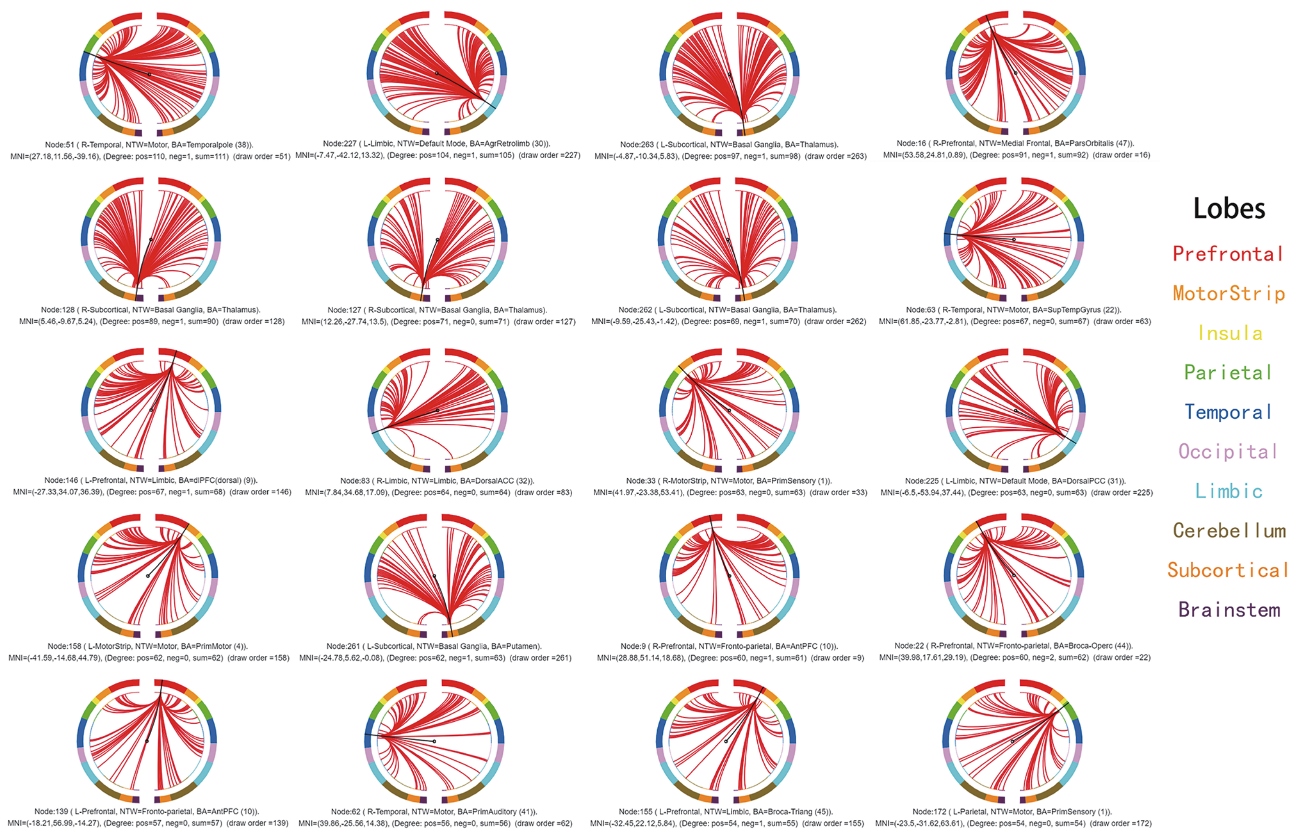

Fig. 3. Connectivity patterns of the top 20 nodes with the most connections. *Abbreviations*: L, left; R, right; BA, Brodmann area; MNI, Montreal Neurological Institute.

Control analyses

After controlling for the potential confounding variables of head motion, sex and age, the predictive models remained statistically significant (Table 4).

External generalizability on dataset 2 ($N = 71$)

We assessed whether the predictive model generated by the discovery dataset could be generalized to an independent and external sample. In the independent external dataset, we observed a significant prediction of TFS for the positive ($r = 0.262$, $P = 0.025$, $MSE = 28.33$; Figure 1C), negative ($r = 0.313$, $P = 0.008$, $MSE = 15.71$) and combined networks ($r = 0.393$, $P = 0.001$, $MSE = 23.69$). These results suggest that the set of edges and parameters identified in the discovery dataset is especially robust in predicting individual differences in trait forgiveness.

Table 3. Results of different CV schemes at the threshold of 0.05

	Positive network			Negative network			Combined network	
	r	*P*	MSE	*r*	*P*	MSE	*r*	*P*
LOOCV								
	0.230	0.021	31.56	0.260	0.009	29.81	0.312	0.002
K-fold								
2	0.222	0.040	36.84	0.023	0.357	40.99	0.194	0.145
5	0.225	0.030	33.10	0.209	0.084	33.69	0.296	0.006
10	0.226	0.026	32.32	0.245	0.018	30.07	0.305	0.002
20	0.229	0.023	31.72	0.252	0.012	29.65	0.309	0.002

Abbreviations: *r*, Pearson correlation coefficient in network; *P*, probability value in network.
Note: this table shows *r*, *P* and MSE in positive, negative and combined networks across different CV schemes.

Table 4. Results of control analyses

	Positive network			Negative network			Combined network	
Control variable	*r*	*P*	MSE	*r*	*P*	MSE	*r*	*P*
Age	0.224	0.025	31.35	0.254	0.011	29.97	0.306	0.002
Sex	0.231	0.021	31.46	0.259	0.009	29.81	0.312	0.002
FD	0.224	0.025	32.52	0.226	0.024	30.76	0.306	0.002
Scrubbing	0.231	0.021	31.62	0.241	0.016	30.21	0.298	0.003

Abbreviations: *r*, Pearson correlation coefficient in network; *P*, probability value in network.

Discussion

In this study, we aimed to predict trait forgiveness in healthy participants using whole-brain rsFC and connectome-based predictive models with different CV and control analyses. The results showed that individual differences in trait forgiveness could be robustly predicted by multiple brain systems, including DMN, FPN, motor network and limbic network, supporting that trait forgiveness is highly associated with multiple brain regions. These findings firstly reveal that intrinsic functional connectivity across multiple neural systems contributes to the prediction of individual differences in trait forgiveness. Specifically, interindividual variations in trait forgiveness were primarily accounted for by intrinsic connectivity within the limbic cortex and PFC, and connectivity with other networks, particularly the temporal and subcortical structures, which are neural underpinnings of cognitive control and perspective taking involved in the process of forgiveness.

Among the top 20 key nodes, the DMN (including the retrosplenial cortex, precuneus and the PCC) contributed the most to positive networks. Recent studies have suggested that the DMN is associated with the experience of the sense of self, and this reduced focus on perceived wrong with the self is the possible neuropsychological foundation of character traits (i.e. forgiveness) (Carhart-Harris, 2018; Johnstone *et al.*, 2021). The retrosplenial cortex, which is involved in emotion and episodic memory, was observed to be a critical region contributing to the predictive model. Retrosplenial cortex activation is associated with negative self-referential scenarios, such as guilt, shame and empathy, which have been identified as moral emotions that may inform forgiveness (Tangney, 1999). In line with this study, guilt proneness is highly correlated with forgiveness, and individuals who tend to score high on perspective taking also tend to score high on guilt (Konstam, 2001). Although no direct link between the retrosplenial cortex and forgiveness has yet been discovered, the potential relationship between the retrosplenial cortex and perspective taking deserves more attention in future studies. Based on this result, we can hypothesize that the retrosplenial cortex plays a critical role in forgiveness mediated by guilt and empathy.

The other key nodes, including the precuneus and PCC in the DMN, were also associated with perspective taking. The PCC supports internally directed thoughts, and the precuneus is related to episodic memory retrieval, self-related mental representations and first-person perspective taking (Cavanna and Trimble, 2006). From the perspective taking aspect, a greater disposition in perspective taking has been associated with a lower incidence of punishment behavior and a higher incidence of forgiveness toward transgressors (Will *et al.*, 2015). Consistently, the retrosplenial cortex, the precuneus, the medial frontal gyrus, the posterior cingulate, the superior temporal sulcus and the inferior parietal lobe constitute the 'moral brain' (Greene and Haidt, 2002), which also implies that forgiveness is a complex social process involving moral judgment requiring the synergy of multiple brain regions. The temporal pole was observed as the highest degree node of the predictive model of forgiveness in light of perspective taking. Emotional processes implemented in the temporal pole are recruited during a successful understanding of another person's mental state (Jimura *et al.*, 2010). Studies have reported the activation of the temporal pole while inferring the emotional state of others (Farrow *et al.*, 2001; Völlm *et al.*, 2006). Compared to the neutral condition, Michl *et al.* (2014) found additional activation in the left superior temporal gyrus, which is a region relevant for perspective taking. Increased functional connectivity of these brain areas, which are highly associated with perspective taking, indicates that perspective-taking ability is necessary for forgiveness, and individual differences in this ability determine people's different levels of forgiveness.

We also revealed that the dlPFC and dACC were key nodes in the prediction of trait forgiveness. The dACC signals internal conflict when one acts in a prosocial manner toward wrongdoers (Moor *et al.*, 2012). Dorsolateral and posterior portions of the PFC support explicit reappraisal of situations reflecting a more general, indirect mechanism to alter emotional associations.

Further, enhanced recruitment of cognitive control is also pivotal when dealing with a transgression and experiencing conflicting desires (e.g. emotional 'punish' *vs* cognitive 'forgive') (Fourie *et al.*, 2020). Because forgiving others is a process involving cognitive, emotional and behavioral changes, prefrontal cognitive control areas are crucial in countering one's own response tendencies and using cognitive strategies to regulate emotions. For example, the lateral prefrontal areas are activated when they reappraise an emotive situation in a positive manner (Drabant *et al.*, 2009) and when they regulate a strong negative affect (Sebastian *et al.*, 2011). A recent study (Maier *et al.*, 2018) provided direct evidence supporting the importance of cognitive control in forgiveness decisions, in which cognitive control was manipulated in real time through inhibitory continuous theta burst stimulation (cTBS) of the dlPFC. Participants who received cTBS displayed significantly more revenge than forgiveness in a dictator game against previously unfair opponents. Consequently, in the present study, altered connectivity in the dlPFC and dACC was identified as potentially underlying or reflecting the variety of cognitive controls, which brings about differences in individuals' personality forgiveness.

In the present study, the anterior PFC and precuneus were cooperatively used to predict the variation in trait forgiveness, which is consistent with a previous study wherein the anterior PFC, superior temporal sulcus and precuneus showed increased coupling in a functional connectivity analysis (Moll *et al.*, 2008). People have the ability to evaluate future outcomes and consider altruistic actions, such as forgiveness (Moll *et al.*, 2005). This inherent difference may be associated with the anterior PFC and limbic regions, which represent social–emotional events linked to 'moral sensitivity' (Moll *et al.*, 2008).

We also found that the areas of the thalamus comprised high-degree nodes in the positive network. The thalamus is considered to provide a greater and more complex contribution toward cognition rather than simply serving as a relay that transfers information (Wolff and Vann, 2019). The thalamus has not previously been associated with the process of forgiveness; future research can further explore its more profound function in cognition and forgiveness.

Taken together, these findings provide evidence that forgiveness is a complex social-cognitive process that requires the coordination of many different social-cognitive abilities, for which a whole-brain functional connectivity approach provides a more comprehensive measure. Consistent with previous studies, the functional connectomes of some networks involving cognitive control and perspective taking are significant predictors of an individual's trait forgiveness among the general population. Some brain areas, such as the dlPFC, dACC and precuneus, have been widely associated with the forgiveness process (Ridderinkhof *et al.*, 2004; Li and Lu, 2017; Maier *et al.*, 2018, 2021). Nevertheless, most of these findings are based on well-controlled designs, which allow flexibility at the analytical level to investigate variations in task-induced processes and representations (Tibon *et al.*, 2022). The whole-brain rsFC networks used in the present study could explore individual differences in forgiveness, which is not related to a highly specific process. Furthermore, rsFC networks can eliminate the risk of overrepresentation of certain ROIs (Sprooten *et al.*, 2017). Thus, our findings provide complementary evidence for previous studies that general social-cognitive abilities, such as cognitive control, perspective taking, social evaluation and moral judgment, regardless of the situation, play a significant role in forgiveness.

The present study had several limitations. Although we tested the stability of the prediction results using various approaches, caution must be exercised when interpreting our findings. First, following the advice that a dataset of over 100 individuals should be used for the predictive modeling (Scheinost *et al.*, 2019), but in comparison with other large, open and shared datasets (Human Connectome Project and The Adolescent Brain Cognitive Development), our findings were based on a relatively small sample. Although we observed significant results in the independent and external datasets, the predictive accuracy might have been overestimated and the generalizability of our findings requires further validation using an independent and larger sample. Second, both positive and negative networks exhibited good predictive models, and the negative network resulted in greater prediction accuracy. However, few nodes in the negative network have high degrees, and most feature edges are in the positive network; therefore, the higher accuracy in negative networks might have resulted from underfitting. In particular, combining the two networks provided a much better predictive model, indicating that the negative network also provides information for prediction. Previous studies have suggested that positive and negative networks represent different functions and are disproportionately located in different functional networks (Finn *et al.*, 2015; Shen *et al.*, 2017; Beaty *et al.*, 2018; Jiang *et al.*, 2018; Feng *et al.*, 2019). In this way, the predictive model should be explained based on both positive and negative networks to obtain a comprehensive understanding of the association between the brain and behavior and enhance the model's interpretability. Third, we observed high-level nodes that had not been presented in previous forgiveness studies. This might have been due to most previous studies being small-scale, task-based studies with well-controlled designs. Although this study was based on resting-state fMRI, the connections or activity of some brain regions may only be observed in some offensive situations or in response to external stimuli. Moreover, the extent to which brain functional connectivity reflects transient states *vs* stable traits remains unknown (Suo *et al.*, 2022).

Despite these limitations, to our knowledge, this study is the first to demonstrate that the functional connectivity of distributed networks effectively predicts trait forgiveness at the individual level. Notably, the nodes and edges of the predictive network are frequently implicated in cognitive control, perspective taking and moral judgment, which are strongly associated with forgiveness and are required for developing and maintaining social connections. The current study's findings may have important implications for characterizing the neural mechanisms of forgiveness.

Data availability

Scripts to run CPM analyses can be found at https://www.nitrc.org/projects/bioimagesuite/. The processed data are available on request from the corresponding author.

Authors' contributions

J.Y.L. and H.J.L. designed the study. J.Q. collected the experimental data. J.Y.L. performed the experiment and analyzed the data. J.Y.L. and H.J.L. wrote the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (31700995) and the Humanity and Social

Science Youth Foundation of Ministry of Education of China (17YJC190011).

Conflict of interest

The authors declared that they had no conflict of interest with respect to their authorship or the publication of this article.

References

Beaty, R.E., Kenett, Y.N., Christensen, A.P., *et al.* (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences*, **115**, 1087–92.

Berry, J.W., Worthington, E.L., O'Connor, L.E., *et al.* (2005). Forgivingness, vengeful rumination, and affective traits. *Journal of Personality*, **73**, 183–226.

Brose, L.A., Rye, M.S., Lutz-Zois, C., *et al.* (2005). Forgiveness and personality traits. *Personality and Individual Differences*, **39**, 35–46.

Burnette, J.L., McCullough, M.E., van Tongeren, D.R., *et al.* (2012). Forgiveness results from integrating information about relationship value and exploitation risk. *Personality & Social Psychology Bulletin*, **38**, 345–56.

Cai, H., Chen, J., Liu, S., *et al.* (2020). Brain functional connectome-based prediction of individual decision impulsivity. *Cortex*, **125**, 288–98.

Carhart-Harris, R.L. (2018). The entropic brain - revisited. *Neuropharmacology*, **142**, 167–78.

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

Drabant, E.M., McRae, K., Manuck, S.B., *et al.* (2009). Individual differences in typical reappraisal use predict amygdala and prefrontal responses. *Biological Psychiatry*, **65**, 367–73.

Eickhoff, S.B., Langner, R. (2019). Neuroimaging-based prediction of mental traits: road to utopia or orwell? *PLOS Biology*, **17**(11), e3000497.

Farrow, T., Zheng, Y., Wilkinson, I., *et al.* (2001). Investigating the functional anatomy of empathy and forgiveness. *NeuroReport*, **12**, 2433–8.

Farrow, T.F.D., Hunter, M.D., Wilkinson, I.D., *et al.* (2005). Quantifiable change in functional brain response to empathic and forgivability judgments with resolution of posttraumatic stress disorder. *Psychiatry Research: Neuroimaging*, **140**, 45–53.

Fatfouta, R., Meshi, D., Merkl, A., *et al.* (2016). Accepting unfairness by a significant other is associated with reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex. *Social Neuroscience*, **13**, 61–73.

Feng, C., Wang, L., Li, T., *et al.* (2019). Connectome-based individualized prediction of loneliness. *Social Cognitive and Affective Neuroscience*, **14**, 353–65.

Finn, E.S., Shen, X., Scheinost, D., *et al.* (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, **18**, 1664–71.

Fourie, M.M., Hortensius, R., Decety, J. (2020). Parsing the components of forgiveness: psychological and neural mechanisms. *Neuroscience and Biobehavioral Reviews*, **112**, 437–51.

Gençoğlu, C., Şahin, E., Topkaya, N. (2018). General self-efficacy and forgiveness of self, others, and situations as predictors of depression, anxiety, and stress in university students. *Educational Sciences: Theory & Practice*, **18**, 605–26.

Greene, J., Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, **6**, 517–23.

Griffin, B.J., Worthington, E.L., Lavelock, C.R., *et al.* (2015). Forgiveness and mental health. In: Toussaint, L., Worthington, E., Williams, D.R., editors. *Forgiveness and Health: Scientific Evidence and Theories Relating Forgiveness to Better Health*, Dordrecht: Springer Netherlands. 77–90.

Jenkinson, M., Bannister, P., Brady, M., *et al.* (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, **17**, 825–41.

Jiang, R., Calhoun, V.D., Zuo, N., *et al.* (2018). Connectome-based individualized prediction of temperament trait scores. *NeuroImage*, **183**, 366–74.

Jimura, K., Konishi, S., Asari, T., *et al.* (2010). Temporal pole activity during understanding other persons' mental states correlates with neuroticism trait. *Brain Research*, **1328**, 104–12.

Johnstone, B., Cohen, D., Dennison, A. (2021). The integration of sensations and mental experiences into a unified experience: a neuropsychological model for the "sense of self". *Neuropsychologia*, **159**, 107939.

Kohavi, R. (1995). The power of decision tables. In: Lavrac, N., Wrobel, S., editors. *Machine Learning: ECML-95*. *Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer. 174–89.

Konstam, V. (2001). Toward forgiveness: the role of shame, guilt anger, and empathy. 14. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2161-007X.2001.tb00204.x.

Lawler-Row, K.A., Hyatt-Edwards, L., Wuensch, K.L., *et al.* (2011). Forgiveness and health: the role of attachment. *Personal Relationships*, **18**, 170–83.

Li, H., Lu, J. (2017). The neural association between tendency to forgive and spontaneous brain activity in healthy young adults. *Frontiers in Human Neuroscience*, **11**, 561.

Maier, M.J., Rosenbaum, D., Haeussinger, F.B. (2018). Forgiveness and cognitive control - provoking revenge via theta-burst-stimulation of the DLPFC. *NeuroImage*, **183**, 769–75.

Maier, M.J., Rosenbaum, D., Haeussinger, F.B., *et al.* (2019). Disinhibited revenge – an fNIRS study on forgiveness and cognitive control. *Frontiers in Behavioral Neuroscience*, **13**, 223.

Maier, M.J., Rosenbaum, D., Brüne, M., *et al.* (2021). The impact of TMS-enhanced cognitive control on forgiveness processes. *Brain and Behavior*, **11**, e02131.

McCullough, M.E., Kurzban, R., Tabak, B.A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, **36**, 1–15.

Michl, P., Meindl, T., Meister, F., *et al.* (2014). Neurobiological underpinnings of shame and guilt: a pilot fMRI study. *Social Cognitive and Affective Neuroscience*, **9**, 150–7.

Moll, J., Zahn, R., de Oliveira-souza, R., *et al.* (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, **6**, 799–809.

Moll, J., De Oliveira-Souza, R., Zahn, R. (2008). The neural basis of moral cognition. *Annals of the New York Academy of Sciences*, **1124**, 161–80.

Moor, B., Güroğlu, B., Op de Macks, Z.A., *et al.* (2012). Social exclusion and punishment of excluders: neural correlates and developmental trajectories. *NeuroImage*, **59**, 708–17.

Ohtsubo, Y., Matsunaga, M., Tanaka, H., *et al.* (2018). Costly apologies communicate conciliatory intention: an fMRI study on forgiveness in response to costly apologies. *Evolution and Human Behavior*, **39**, 249–56.

Ren, Z., Daker, R.J., Shi, L., *et al.* (2021). Connectome-based predictive modeling of creativity anxiety. *NeuroImage*, **225**, 117469.

Ricciardi, E., Rota, G., Sani, L., *et al.* (2013). How the brain heals emotional wounds: the functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience*, **7**, 839.

Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., *et al.* (2004). The role of the medial frontal cortex in cognitive control. *Science*, **306**, 443–7.

Rosenberg, M.D., Finn, E.S., Scheinost, D., *et al.* (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, **19**, 165–71.

Rudebeck, P.H., Bannerman, D.M., Rushworth, M.F.S. (2008). The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cognitive, Affective & Behavioral Neuroscience*, **8**, 485–97.

Rutherford, H.J.V., Potenza, M.N., Mayes, L.C., *et al.* (2020). The application of connectome-based predictive modeling to the maternal brain: implications for mother–infant bonding. *Cerebral Cortex*, **30**, 1538–47.

Scheinost, D., Noble, S., Horien, C., *et al.* (2019). Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage*, **193**, 35–45.

Sebastian, C.L., Tan, G.C.Y., Roiser, J.P., *et al.* (2011). Developmental influences on the neural bases of responses to social rejection: implications of social neuroscience for education. *NeuroImage*, **57**, 686–94.

Shen, X., Papademetris, X., Constable, R.T. (2010). Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. *NeuroImage*, **50**, 1027–35.

Shen, X., Tokoglu, F., Papademetris, X., *et al.* (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage*, **82**, 403–15.

Shen, X., Chun, M.M., Papademetris, X., *et al.* (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols*, **12**, 506–18.

Sprooten, E., Rasgon, A., Goodman, M., *et al.* (2017). Addressing reverse inference in psychiatric neuroimaging: meta-analyses of task-related brain activation in common mental disorders. *Human Brain Mapping*, **38**, 1846–64.

Suo, X., Zuo, C., Lan, H., *et al.* (2022). COVID-19 vicarious traumatization links functional connectome to general distress. *NeuroImage*, **255**, 119185.

Tangney, J.P. (1999). The self-conscious emotions: shame, guilt, embarrassment and pride. In: Dalgleis, T., Power, M.J., editors. *Handbook of Cognition and Emotion*, New York, NY: John Wiley & Sons Ltd, 541–68.

Tibon, R., Geerligs, L., Campbell, K. (2022). Bridging the big (data) gap: levels of control in small- and large-scale cognitive neuroscience research. *Trends in Neurosciences* **45**(7), 507–16.

Toussaint, L., Friedman, P. (2008). Forgiveness, gratitude, and well-being: the mediating role of affect and beliefs. *Journal of Happiness Studies*, **10**, 635.

Varoquaux, G., Raamana, P.R., Engemann, D.A., *et al.* (2017). Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *NeuroImage*, **145**, 166–79.

Völlm, B.A., Taylor, A.N.W., Richardson, P., *et al.* (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, **29**, 90–8.

Wang, Z., Goerlich, K.S., Ai, H., *et al.* (2021). Connectome-based predictive modeling of individual anxiety. *Cerebral Cortex*, **31**, 3006–20.

Whelan, R., Garavan, H. (2014). When optimism hurts: inflated predictions in psychiatric neuroimaging. *Biological Psychiatry*, **75**, 746–8.

Will, G.-J., Crone, E.A., Güroğlu, B. (2015). Acting on social exclusion: neural correlates of punishment and forgiveness of excluders. *Social Cognitive and Affective Neuroscience*, **10**, 209–18.

Wolff, M., Vann, S.D. (2019). The cognitive thalamus as a gateway to mental representations. *Journal of Neuroscience*, **39**, 3–14.

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

Yan, C.-G., Craddock, R.C., Zuo, X.-N., *et al.* (2013). Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. *NeuroImage*, **80**, 246–62.

Yang, F.N., Hassanzadeh-Behbahani, S., Bronshteyn, M., *et al.* (2021). Connectome-based prediction of global cognitive performance in people with HIV. *NeuroImage: Clinical*, **30**, 102677.

Yeung, A.W.K., More, C., Wu, J., *et al.* (2022). Reporting details of neuroimaging studies on individual traits prediction: a literature survey. *NeuroImage*, **256**, 119275.