Behavioral/Cognitive

Decomposing Gratitude: Representation and Integration of Cognitive Antecedents of Gratitude in the Brain

Hongbo Yu,[1,2,3]* Xiaoxue Gao,[1,2]* Yuanyuan Zhou,[1,2,4] and Xiaolin Zhou[1,2,5,6,7,8]

[1]Center for Brain and Cognitive Sciences, Peking University, Beijing 100871, China, [2]School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China, [3]Department of Experimental Psychology, University of Oxford, OX1 3UD, Oxford, United Kingdom, [4]School of Economics and Management, Tsinghua University, Beijing 100084, China, [5]Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China, [6]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China, [7]Institute of Psychological and Brain Sciences, Zhejiang Normal University, Zhejiang 321004, China, and [8]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

Gratitude is a typical social-moral emotion that plays a crucial role in maintaining human cooperative interpersonal relationship. Although neural correlates of gratitude have been investigated, the neurocognitive processes that lead to gratitude, namely, the representation and integration of its cognitive antecedents, remain largely unknown. Here, we combined fMRI and a human social interactive task to investigate how benefactor's cost and beneficiary's benefit, two critical antecedents of gratitude, are encoded and integrated in beneficiary's brain, and how the neural processing of gratitude is converted to reciprocity. A coplayer decided whether to help a human participant (either male or female) avoid pain at his/her own monetary cost; the participants could transfer monetary points to the benefactor with the knowledge that the benefactor was unaware of this transfer. By independently manipulating monetary cost and the degree of pain reduction, we could identify the neural signatures of benefactor's cost and recipient's benefit and examine how they were integrated. Recipient's self-benefit was encoded in reward-sensitive regions (e.g., ventral striatum), whereas benefactor-cost was encoded in regions associated with mentalizing (e.g., temporoparietal junction). Gratitude was represented in perigenual anterior cingulate cortex (pgACC), the strength of which correlated with trait gratitude. Dynamic causal modeling showed that the neural signals representing benefactor-cost and self-benefit passed to pgACC via effective connectivities, suggesting an integrative role of pgACC in generating gratitude. Moreover, gyral ACC plays an intermediary role in converting gratitude representation into reciprocal behaviors. Our findings provide a neural mechanistic account of gratitude and its role in social-moral life.

Key words: **cognitive antecedents; dynamic causal modeling; fMRI; gratitude; integration; reciprocity**

Significance Statement

Gratitude plays an integral role in subjective well-being and harmonious interpersonal relationships. However, the neurocognitive processes through which various components and antecedents of gratitude are integrated remain largely unknown. We developed a new interpersonal paradigm to independently and parametrically manipulate two antecedents of gratitude in a helping context, namely, the benefit to beneficiary and the cost to benefactor, to examine their representation and integration in the beneficiary's brain using fMRI. We found the neural encoding of self-benefit and benefactor-cost in reward- and mentalizing-related brain areas, respectively. More importantly, by examining effective connectivity, we showed that these componential signals are passed to perigenual anterior cingulate cortex, which tracks trial-by-trial gratitude levels. Our study thus provides a neural mechanistic account of gratitude.

Introduction

G a - ; . a a . a . a a . a . (M - C . . - a . 2001; Ha . a , 2004; M C . . a . T a , 2004; L ; a ; 2014; K - a . , 2015; Ma . a, 2015) a . . a . a a - a . - . (Ha . ; 2003; A a . Ha . ; 2009). T . . a . a . \ a a a . a - ; . .

Received Oct. 26, 2017; revised April 7, 2018; accepted April 17, 2018.

Author contributions: H.Y. and X.G. edited the paper; Y.Z. wrote the first draft of the paper. H.Y., X.G., Y.Z., and X.Z. designed research; H.Y. and X.G. performed research; H.Y., X.G., and Y.Z. analyzed data; H.Y., X.G., Y.Z., and X.Z. wrote the paper.

This work was supported by the Natural Science Foundation of China Grants 31630034 and J1103602 and the National Basic Research Program of China 973 Program: 2015CB856400. H.Y. was supported by British Academy Newton International Fellowship NF160700. An earlier version of the manuscript and data analysis have been presented to the research groups led by Dr. Paul Glimcher, Dr. Matthew Rushworth, Dr. Molly Crockett, and Dr. Bolton

Materials and Methods

Participants

Experimental design and statistical analyses

Overview.

Randomization and pain titration.

Pain titration.

Chau. The authors have received invaluable comments and suggestions concerning behavioral modeling and imaging analysis from discussions with these groups. We thank Joshua Brown for sharing the activation maps from Kini et al. (2016) for comparison; and Dr. Philip Blue for commenting on an earlier version of the manuscript.

The authors declare no competing financial interests.

*H.Y. and X.G. contributed equally to this work.

Correspondence should be addressed to Dr. Xiaolin Zhou, School of Psychological and Cognitive Sciences Peking University, Beijing, 100871 China. E-mail: xz104@pku.edu.cn.

DOI:10.1523/JNEUROSCI.2944-17.2018

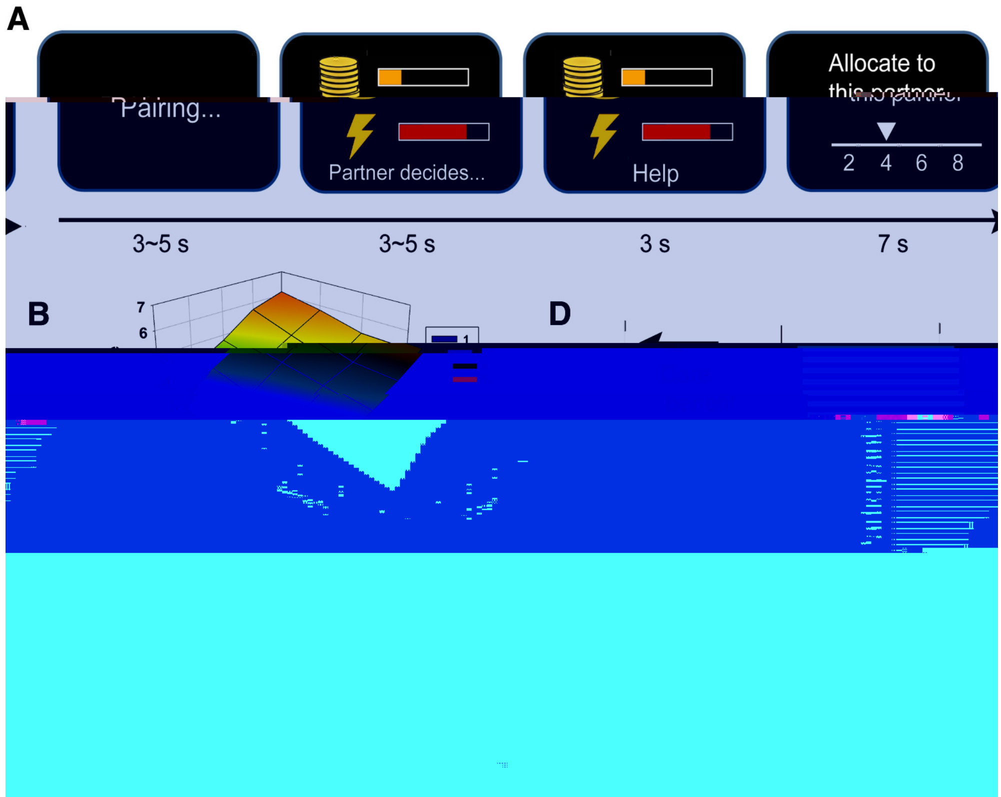

Figure 1. Experiment procedure and behavioral results. ***A***, At the beginning of each trial, the participants were (anonymously and ostensibly) paired with 1 of 3 coplayers. Then the participants saw a pain-money pair and waited for the coplayer's decision. If the coplayer chose Help, then the coplayer lost the corresponding amount of bonus while the participants would be exempted from the pain stimulation on this trial. If the coplayer chose NoHelp, then the coplayer could keep the bonus while the participants had to receive the corresponding pain stimulation. The presentation of the coplayer's decision was defined as the critical events in the fMRI data analysis. At the end of the trial, the participant could allocate 20 Yuan (∼$3 U.S.) between himself/herself and the coplayer, with the knowledge that the coplayer was not aware of this procedure. ***B***, ***C***, Postscan gratitude rating and allocation during scanning (i.e., reciprocity) as a function of self-benefit and benefactor-cost. ***D***, Relative weight of benefactor-cost over self-benefit in gratitude rating.

2 … A … (2, 4, 8, 12 …) …

Pain-money exchange task (behavioral). A … Ea … I … 15 … T … I … 4 … 0.5, 1.0, 1.5, … 2.0 Y … (1 Y … ∼ $0.16 U.S.), … T … F … /a … T … 1.5. F … 0.8, 1.5, 2.3, … 3.0 Y … O … (…), … 0.5, … 0.5. A … 0.1 … (… S … a., 2016). T … U … 18 Y … (∼$2.8 U.S.). B …

Help-receiving task (fMRI). I … BOLD … MRI (F … 1*A*). I … T … (1 … 4, … 5 … (0 … 4, … 0%, 25%, 50%, 75%, … 100% …). T … (H … N H …). I … W … 20 … F … T … MRI … A … 20 Y … (∼$3 U.S.) … T … 2 Y … T … Pa … MRI … T … T … (…). T …

Table 1. Distribution of NoHelp trials in different cost-benefit conditions

No. of NoHelp trials	Cost 0	1	2	3	4
Benefit					
1	0	3	3	5	5
2	0	3	3	4	5
3	0	2	3	3	3
4	0	1	2	3	3

[illegible]

Analysis of the behavioral data. [illegible]

[illegible] $(\beta = 0.88 \pm 0.09$, [illegible] $R^2 = 0.44)$.

MRI data acquisition and preprocessing. [illegible] 3.0 T MR [illegible] (GE MR750) [illegible] (TR = 2000 [illegible]; TE = 30 [illegible]; [illegible] = 90; FOV = 192 [illegible] × 192 [illegible]; [illegible] = 4 [illegible]). [illegible] SPM8 [illegible] MNI [illegible] 8 [illegible] FWHM [illegible] 1/128 H [illegible]

fMRI data analysis. [illegible] GLM1 [illegible] GLM2 [illegible]

R GLM1.

T H > N H GLM2 . A ( ) , *t* ’ (. .,) ( ). P (H > N H) . I , , F GLM , . S *t* . A *p* < 0.001 ( ) *p* < 0.05 (FWE-). T

cost were significant (benefit: 0.30 0.07, t 4.58; cost: 0.80 0.16, t 5.02); the interaction term did not reach significance (0.01 0.05, t 0.26). For allocation (Table 3), the model with the two main effects was the best model. Parameters estimated based on this model showed that both benefit and cost were predictive of allocation (benefit: 0.64 0.10, t 6.11; cost: 0.78 0.11, t 7.43). Both gratitude rating and allocation increased monotonically with benefit and cost (Fig. 1B, C).

To examine whether the participants' allocation was influenced by trial history, namely, trial features (cost, benefit) and the benefactor's decision from the previous trial, we performed a separate regression model for allocation (Help trials alone) with this information included the following:

$$\text{Allocation}_n = {}_0 + {}_1\text{Cost}_n + {}_2\text{Benefit}_n + {}_3\text{Cost}_{n-1} + {}_4\text{Benefit}_{n-1} + {}_5\text{Decision}_{n-1} + {}_6\text{Decision}_{n-1} \times \text{Cost}_{n-1} + {}_7\text{Decision}_{n-1} \times \text{Benefit}_{n-1} \quad (2)$$

We found that the contribution of cost and benefit on the trial remained significant ($_1$ 2.07 0.18, t 11.18; $_2$ 0.55 0.14, t 3.96). Interestingly, the contribution of benefit on the last trial was also significant ($_4$ 0.18 0.08, t 2.26), and it was qualified by a significant interaction with benefactor's decision ($_7$ 0.27 0.10, t 2.65). These results indicate that the participants allocated more on the current trial if the benefit on the last trial was high and the benefactor chose "Help." Benefactor's sacrifice on trial $_{n-1}$ did not influence participants' allocation on trial $_n$, nor did its interaction with benefactor's decision. These findings shed light on how the impacts of different cognitive antecedents on gratitude and reciprocity persist and decay over time. Decisive conclusion in this regard is beyond the scope of the current study because this study was not designed to address the question; thus, it did not balance the distribution of cost, benefit, and benefactor's decision over time.

fMRI results

Neural representations of cost and benefit

Our first aim was to examine how the brain encodes benefit and cost when receiving help. Contrasts corresponding to the main effect of benefactor-cost and self-benefit in the Help conditions were defined based on the regressors in GLM 1 (see Materials and Methods). As we predicted, the main effect of benefactor-cost (Contrast 1) revealed activations in dorsomedial PFC, precuneus, and bilateral TPJs, the regions implicated in empathy and mentalizing (Table 4, Fig. 2A). The main effect of self-benefit (Contrast 2) revealed activations in a network related to value representation, including the ventromedial PFC, bilateral VS, and dorsal striatum (Table 4, Fig. 3A). Regional activation patterns were extracted from our hypothesized ROIs for illustrative purposes (Figs. 2B, 3B). As a comparison, the same set of contrasts defined for the NoHelp trials revealed no suprathreshold activation at the brain areas revealed by the corresponding contrasts in the Help trials (Figs. 2A, 3A). It is worth noting, however, that the null effect of the NoHelp contrasts is not sufficient to demonstrate that the neural processes observed here are specific to receiving help. To demonstrate specificity, one needs to show "separate modifiability" (e.g., Woo et al., 2014) of two constructs (e.g., Help vs NoHelp), which is beyond the scope of the current study.

Because our primary interest here is the neurocognitive processes underlying receiving help and feeling grateful, we included the NoHelp conditions only as fillers. As a result, the distribution of NoHelp trials was not balanced across different levels of cost and benefit, neither was it matched with the Help trials (Table 1). Future studies are needed to reveal the cognitive and affective response to other's withdrawal of help.

Neural representation of gratitude

Directly examining the representation of gratitude required us to have for each participant a trial-by-trial measure of gratitude and perform a parametric regression against brain activity elicited by

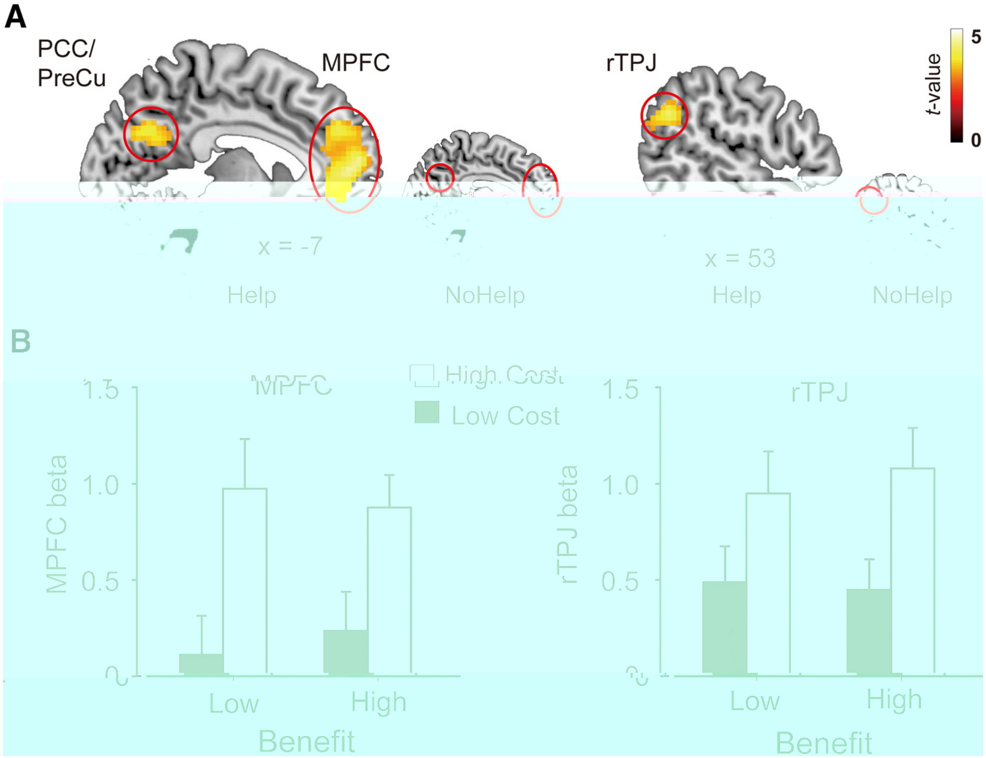

Figure 2. Encoding of benefactor-cost. ***A***, Whole-brain contrast of high versus low cost in Help conditions (larger figure). The same contrast in the NoHelp conditions was inserted for comparison (smaller figure). ***B***, Parameter estimates (β values) corresponding to the four Help conditions were extracted from MPFC and rTPJ for illustrative purposes. Error bars indicate standard error of means.

a a). a ACC a .

I a ACC a a a . W a β a a a a a (C a 4) ACC. I , a a a a a a a (F . 4B, a). M , a a a ACC' a a (F . 4C, a), a a a a ACC (a a , a) a a a a .

Neural integration of cost and benefit

O a a a (. ., VS a TPJ), a a a a a a a . W a a a a a VS a TPJ, a ACC a a a a a . W a a 33 a , a ; a . T 7 a . M a a a a a (F . 5A). Ba a a a a M Fa 1 a a a (0.33; F . 5B). I a , VS a TPJ a a ACC. T a a a a a a a a a (. ., ACC). T a a Ba a a a M Fa 1 a a VS

a TPJ ACC a (Ta 5). M , H C _L B , a a TPJ ACC (F . 5D). I a ; a VS ACC. O a , DCM a a a a a . T a a a a a a a a a a , a a a (F . 1D). A a a a a a a a a a , a a a a a a a a a a a a .

A a F 5B, M Fa 2 a a a a a (0.28). G a Fa 1 a Fa 2 a a TPJ a VS ACC, a a a a a a a a a a a a a a . T a a Fa 2 a VS TPJ. T a a a a a a a / ; a ROI- a a a a a VS a TPJ. S a , VS a a a a TPJ (Ma a a M). T a VS a TPJ a ($F_{(1,30)} = 0.18, p = 0.68$) ($F_{(1,30)} =$

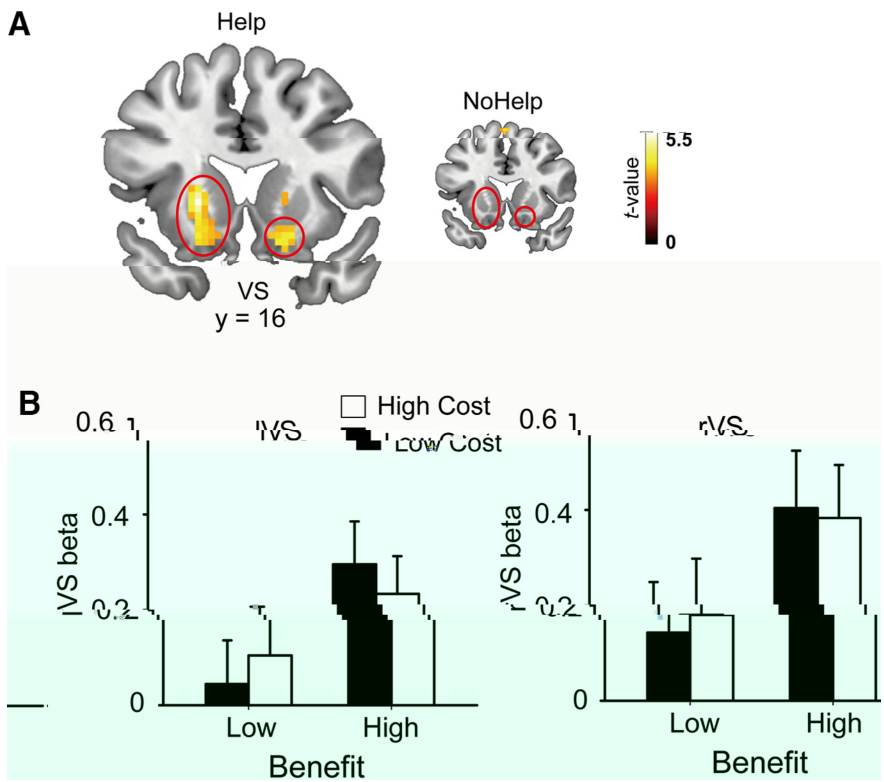

Figure 3. Encoding of self-benefit. ***A***, Whole-brain contrast of high versus low benefit in Help conditions (larger figure). The same contrast in the NoHelp conditions was inserted for comparison (smaller figure). ***B***, Parameter estimates (β values) corresponding to the four Help conditions were extracted from left VS and rVS for illustrative purposes. Error bars indicate standard error of means.

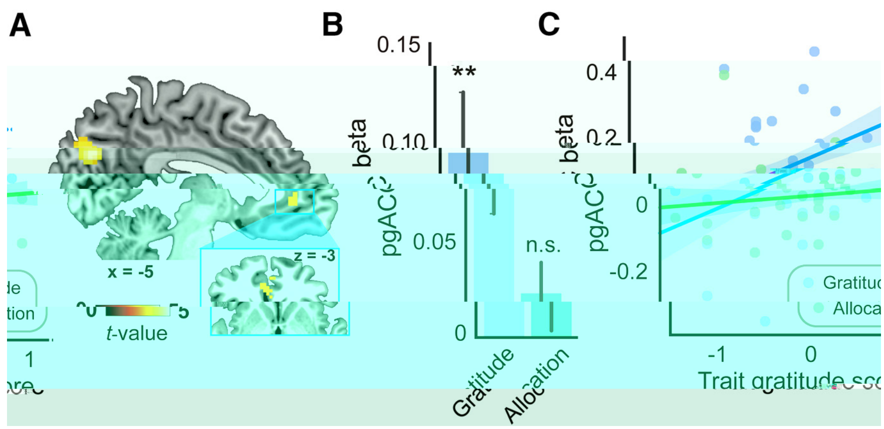

Figure 4. Representation of gratitude. ***A***, Whole-brain parametric contrast of constructed gratitude. ***B***, pgACC responses to constructed gratitude (blue) and allocation (green). ***C***, Relation between trait gratitude score and pgACC responses to constructed gratitude (blue) and allocation (green). $^{**}p < .005$. Error bars indicate standard error of means.

0.00, $p = 0.97$) (Ta 6), … VS a TPJ … BOLD … (A … a., 2015).

From gratitude to reciprocity

N … (Y … a., 2017), … S …

… ($r = 0.47$, $p = 0.029$; F . 6A). A … ($t = 11.68$; F . 6B). H … F 6C, … T … H > N H … GLM 2 (F . 6D). T … (Ba …, 1987; F a S …, 2006). O … P … (C a … a., 2011). W … ACC … C a … a. (2011). W … ([−5, 23, 28], $t = 3.03$, $p_{FWE} = 0.041$, …) … O … T … A a Ra a (2014). T … ACC … W … ([−3, 20, 22], $t = 3.19$, $p_{FWE} = 0.030$, …) … T … I …

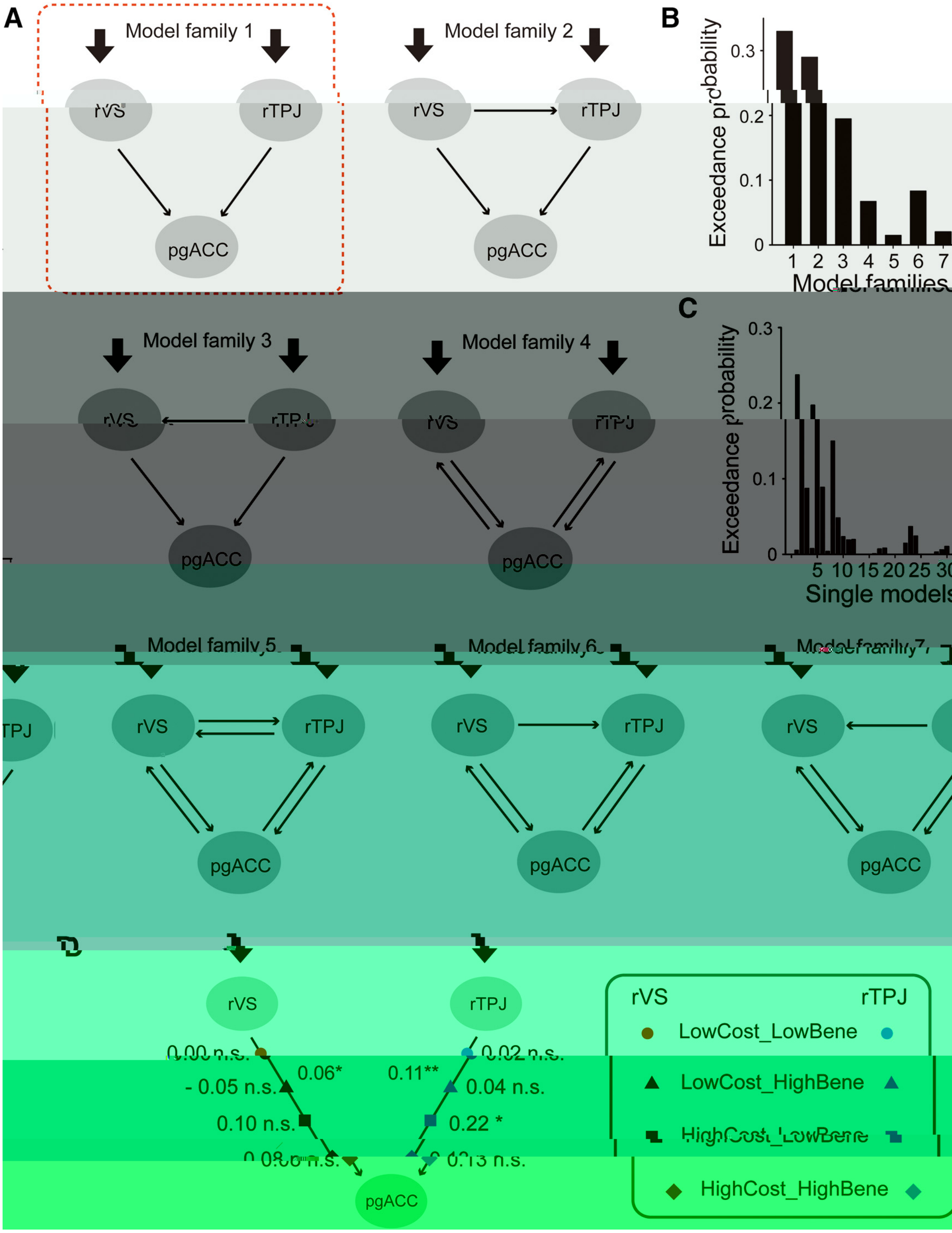

Figure 5. Results of effective connectivity (DCM) analysis. ***A***, Thirty-three individual models, grouped into 7 model families, were constructed and compared using Bayesian Model Comparison. The exceedance probability of each family (***B***) and each individual model (***C***) are shown. Model Family 1, enclosed in the red square, has the highest exceedance probability. ***D***, Strength of intrinsic and modulatory connectivities estimated based on the winning family. $^*p < 0.05$, $^{**}p < .005$.

Discussion

Table 5. Model parameters estimated based on Model Family 1

Parameter	Mean ± SD
Intrinsic connectivity	
VS → pgACC	0.06 ± 0.14*
rTPJ → pgACC	0.11 ± 0.18**
Modulation on VS → pgACC	
LowCost_LowBene	0.00 ± 0.06
LowCost_HighBene	−0.05 ± 0.42
HighCost_LowBene	0.10 ± 0.30
HighCost_HighBene	0.08 ± 0.67
Modulation on rTPJ → pgACC	
LowCost_LowBene	0.02 ± 0.06
LowCost_HighBene	0.04 ± 0.39
HighCost_LowBene	0.22 ± 0.59*
HighCost_HighBene	0.13 ± 0.60
Driving input to VS	
Help decision	0.04 ± 0.08*
Driving input to TPJ	
Help decision	0.06 ± 0.15*

VS, ventral striatum; TPJ, temporoparietal junction; pgACC, perigenual anterior cingulate cortex. *$p < 0.05$; **$p < 0.005$.

Table 6. Functional connectivity (PPI) between rVS and rTPJ

Condition	Connectivity (mean ± SD)
LowCost_LowBene	1.51 ± 4.10
LowCost_HighBene	1.17 ± 2.69
HighCost_LowBene	1.44 ± 3.25
HighCost_HighBene	1.83 ± 4.95

[illegible]

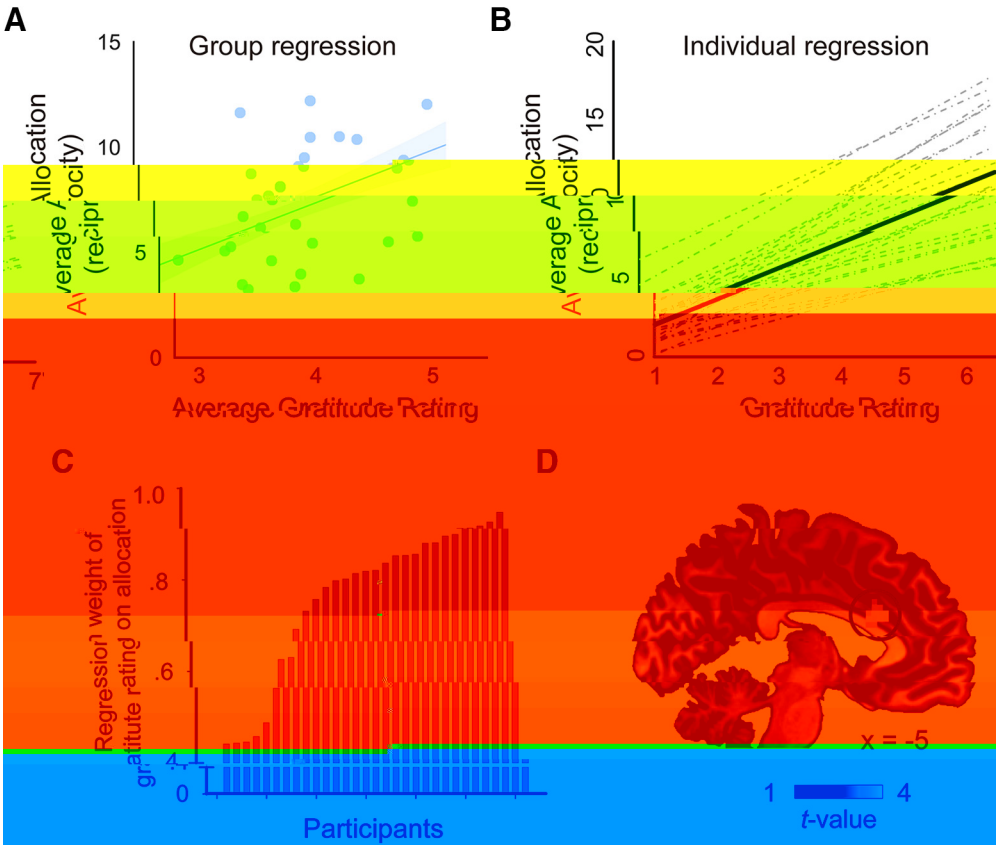

Figure 6. From gratitude to reciprocity. ***A***, Average gratitude rating in the postscan gratitude recall of an individual participant predicts average monetary allocation (i.e., reciprocity) of that participant. ***B***, Within each individual, variation in gratitude ratings predicts variation in allocation. The correlation reported here is the correlation between the postscan gratitude ratings in each of the 20 Help conditions and the average amount of allocation in the 20 Help conditions. Each dotted line indicates the regression line of a single participant, Solid line indicates the group effect. ***C***, Individual differences in the exchange rate between gratitude and reciprocity (i.e., the slopes of the dotted lines in ***B***). ***D***, Neural correlates of individual differences in the exchange rate. This map is thresholded with $p < 0.005$ for illustrative purposes.

(F Ba , 2003) a . I (T a., 1968), , (F a., 2015; K a., 2016; Ka a., 2017; Y a., 2017), . T (., S a, 1935; B , 1975; Ca , 1988; M C , 1993; H a , 2012; G a., 2013). H , , . O . W a ACC (F . 6*D*). T a (A a Ra a , 2014). I , . A , (F a., 1982; Na a F , 1986; Wa a., 2006; Ma a, 2016). I , (F a S , 2006). T a ACC -a (C a a., 2011) . F I , a (E a H a a, 2016), , , a (V a., 2009; B a , 2013). I . B MRI, . W a ACC a , (., a) a (., a). O a W a, , a ,

References

A SB (2012) F , , a : a . S P a P C a 6:455 469. C R

A SB, Ha J (2009) W a : ' a , a , a a a . J P P 4:105 127. C R M

A SB, Ha J, Ga SL (2008) B : a a a . E 8:425 429. C R M

A M, B D, La AR, F PT, S a L, E SB (2015) D a a a a a a . B a S F 220:1031 1049. C R M

A MA, Ra a N (2014) T a a a a ' a . J N 34:6190 6200. C R M

Ba DJ, L R, S C, T HJ (2013) Ra a : a a . J M La 68:255 278. C R M

Ba MY, D S D (2006) G a a a : . P S 17:319 325. C R M

Ba a O, M G JT, Ka JW (2013) T a a : a a a-a a BOLD MRI a a a . N a 76:412 427. C R M

Ba CD (1987) P a a : a ? I : A a a , V 20, 65 122. Sa D , CA: A a .

B FR (1975) G a . E 85: 298 309. C R

B a RJ (2013) T a a . Na R N 14:786 799. C R M

Ca C (1988) G a a a . A P Q 25:115 127.

C a LJ, S A, D M, Sa AG (2011) T a a a , a , a a a . N 70: 560 572. C R M

C MJ, S JZ, K -N Z, Da a P, D a RJ (2017) M a a a a a . Na N 20: 879 885. C R M

Da B, H Y, K F, W B (2017) O a a a a - a a : a MRI . S R 7:43024. C R M

D S D, Ba MY, Ba a J, W a LA, D L (2010) G a a a : - a a . E 10:289 293. C R M

D P, R J, K E, D JC (2018) T - a a a - a a . C C 28:585 601. C R M

E J, H a P (2016) T a . N Y , NY: S .

E PC, S KR (2003) A a a . I : Ha a (Da RJ, S KR, G HH,), 572 595. O , UK: O UP.

E RA, M C ME (2003) C : a a a a a - a . J P S P 84:377 389. C R M

F E, S KM (2006) T a , a a - : a a . I : Ha , a a (K S, Y JM,), 615 691. A a , N a : E .

F MJ, Ba JA (2003) T a a a . I : T a a : a a - (M J, K a K,), 169 188. Ma a , NJ: E a .

F JD, Na A, W -A a S (1982) R a a . P B 91:27 54. C R

F GR, Ka a J, Da a H, Da a A (2015) N a a a - . F P 6:1491. C R M

F BL (2004) G a , , a a . I : T a (E RA, M C ME,), 145 166. N Y , NY: O UP.

F a NH (1993) T a a a a . C E 7:357 387. C R

F KJ, B C, F GR, M J, R E, D a RJ (1997) P a a a a . N a 6:218 229. C R M

F KJ, Ha L, P W (2003) D a a a . N a 19:1273 1302. C R M

G L, M a B, K a K (2013) R a a . J Va I 47:285 317. C R

Ha J (2003) T a . I : Ha a (Da RJ, S KR, G HH,), 852 870. O , UK: O UP.

Ha a EJ (2004) G a a . I : T a (E RA, M C ME,), 19 36. N Y , NY: O UP.

H a B (2012) B a a : , , a . J P 109:391 411. C R

H J, L Y, Y Y, B PR, Y H, Z X (2017) H - a - a a a a ? N a 157:598 611. C R M

H L, Z a L, C R, Y H, L H, M a A (2015) T a a a a a a a - a . H B a Ma 36:4346 4360. C R M

I K, T a TD, H a a M, Ka R (2002) P a a Aδ a- a a . Pa 96:247 252. C R M

Ka CM, M WE 3 , Ma U (2017) T a a a : a a MRI a a a . F H N 11:599. C R M

K SJ, K S, W N, F KJ (2007) D a a : a MRI. N a 34:1487 1496. C R M

K P, W J, M I S, Ga a a N, B JW (2016) T a a a . N a 128:1 10. C R M

K N, W a MK, B TE, B a ED, Ma RB, R MF (2016) Va , a , a a a . Na N 19:1280 1285. C R M

K a K (2015) A A a a . T 34:499 511. C R

La a RS, S CA (1988) K a a a a . C E 2:281 300. C R

L a PJ (2014) G a : a a . Wa , TX: Ba UP.

L a S, Fa S (2010) C a a : a a . T a Oa , CA: Sa .

L SL, O'D JP, Ra A (2013) S a a a PFC a a a a a a . J N 33: 8729 8741. C R M

L KA, Ba LF (2012) A a a a : . T C S 16:533 540. C R M

Ma a T (2015) G a . I T Sa E a P (S 2015 E), E a N. Za a ( ). A a a a :// a . a /a / 2015/ / a .

Ma a T (2016) N a a . J Va I 50:129. C R

M C T (1993) G a . P a a, PA: T UP.

M C ME, T a J (2004) Pa ? T a a . I : T a (E RA, M C ME,), 123 141. N Y , NY: O UP.

M C ME, K a SD, E RA, La DB (2001) I a a a ? P B 127:249 266. C R M

M C ME, E RA, T a JA (2002) T a : a a a a . J P S P 82:112 127. C R M

M a B, G L, K a K (2017) A a a a : a a a . P I D 107:179 189. C R

Na A, F JD (1986) T a a a : a a

a a . I : A a , V 19 (B L,), 81 122. Sa D , CA: A a .
Na T, Wa a J, Ta M (2005) G a Ja a a T a a . J C C P 36:247 263. C R
P WD, S a KE, M A, F KJ (2004) C a a a a . N a 22:1157 1172. C R M
P WD, S a KE, Da a J, R a MJ, F KJ, S TM, L AP (2010) C a a a a a a . PL S C B 6: 1000709. C R M
R CC, F E (2014) T a a a a a . Na R N 15:549 562. C R M
S a (1935) D B . I : S a a a (Ba JW, a a), V III. Ca , MA: L C a a L a , Ha a .
S G, Ca X, S a B, D JC (2013) P a a a a :a a a a-a a a a . N B a R 37:681 696. C R M
S B, Y Y, Wa J, Z X, M C. SM, L J (2016) H