Supporting information for

The hidden cost of receiving help:

A theory of indebtedness

Xiaoxue Gao, Eshin Jolly, Hongbo Yu, Huiying Liu, Xiaolin Zhou*, Luke J. Chang*

*Correspondence to:

Xiaolin Zhou (xz104@pku.edu.cn) and Luke J. Chang (luke.j.chang@dartmouth.edu)

This PDF file includes:

SI Methods

Participants

For Study 1 (large-scale experience sampling), 1,808 graduate and undergraduate students were recruited from Zhengzhou University to complete an online questionnaire regarding the definition of indebtedness and daily-life events of receiving and rejecting help. One hundred and eighty-nine participants were excluded from data analysis because of filling in irrelative information in the essay question, leaving 1,619 participants (812 females, 18.9 ± 2.0 (SD) years) for data analysis. While 98.7% participants reported the events of receiving help, 24.4% participants reported the events of rejecting help within the past one year, which resulted in 1,991 reported effective daily events. To extract the words related to emotions and feelings in the definition of indebtedness, 80 additional graduate and undergraduate students (45 females, 22.6 ± 2.58 years) were recruited from different universities in Beijing to complete the word classification task.

For Study 2a (behavioral study), 58 graduate and undergraduate students were recruited from Zhengzhou University, and 7 participants were excluded due to equipment malfunction, leaving 51 participants (33 females, 19.9 ± 1.6 years) for data analysis. *For Study 2b (behavioral study)*, 60 graduate and undergraduate students were recruited from Zhengzhou University, and 3 participants were excluded due to failing to respond in more than 10 trials, leaving 57 participants (45 females, 20.1 ± 1.8 years) for data analyses.

For Study 3 (fMRI study), 57 right-handed healthy graduate and undergraduate students from Beijing took part in the fMRI scanning. Four participants with excessive head movements (>2mm) were excluded, leaving 53 participants (29 females, 20.9 ± 2.3 years) for data analysis.

None of the participants in the three experiments reported any history of psychiatric, neurological, or cognitive disorders. All the participants were tested only once (i.e.,

for one experiment). Informed written consent was obtained from each participant before each experiment. All the studies were carried out in accordance with the Declaration of Helsinki and were approved by the Ethics Committee of the School of Psychological and Cognitive Sciences, Peking University.

Procedure

Study 1 (large-scale experience sampling)

In this experiment, participants filled in the questionnaire on the Questionnaire Star platform (https://www.wjx.cn/) through their mobile phones. The detailed questionnaire is shown in Appendices S1. Before the formal questionnaire, the participants first pressed the button to confirm that they had at least 15 minutes to fill in the questionnaire and answered each question sincerely.

The questionnaire consists of two parts. In the first part, each participant was asked to recall a daily event in which they **received help** from others, and to answer the corresponding questions. The event should happen within the past year and could be clearly recalled. In the second part, each participant was asked to recall a daily event in which they **rejected help** from others, and to answer the corresponding questions. Again, the event should occur within one year and could be clearly recalled. If the participant did not report an event in which they accepted or rejected another person's help within the past year, questions in the corresponding part were automatically skipped. Appraisal theory suggests that emotions are adaptive responses that: (1) are elicited based on how an agent evaluates its situation (i.e., appraisals) [1-3], and (2) function to motivate the agent's adaptive behavioral responses [4]. Under this framework, questions in each part were designed according to the four levels of the appraisal theory (contextual information, appraisal, emotion and behavior). Specifically, for events in which participants received help (Part 1), questions for contextual information mainly included the self-reported benefactor's cost, the participant's benefit from the help, and the social distance between the benefactor and the participant before this event happened. Questions for appraisals mainly included:

"To what extent do you think the benefactor cared about your welfare? (i.e., perceived care)" and "To what extent do you think the benefactor expected you to repay? (i.e., second-order belief)". Questions for emotions mainly included the ratings for gratitude, indebtedness, guilt and the sense of obligation. Questions for behaviors included: "To what extent did you think you needed to reciprocate?" "To what extent are you willing to reciprocate this benefactor?" "To what extent do you want to accept/reject this offer? " and " To what extent are you willing to interact with this benefactor in the future?" For events in which participants rejected help (Part 2), questions were the same as Part1 except that participants were asked to imagine what they would feel or behave if they accepted this help. See Appendices S1 for questionnaire details. After answering the questions above, participants were asked to briefly describe each event in text without limitation for number of words.

In order to explore how Chinese people define indebtedness, after the event recall questionnaire, participants answered the following three questions about the definition of gratitude and indebtedness:

• In the context of helping and receiving help, what is your definition of gratitude?

• In the context of helping and receiving help, what is your definition of indebtedness?

• In daily life, what do you think is/ are the source(s) of indebtedness? (Single choice, the order of the first two options was counterbalanced over participants)

- Negative feeling for harming the benefactor/for cost that the benefactor has paid for helping you

- Negative feeling for the pressure to repay caused by other's ulterior intentions (e.g., Expectation for repayment)

- Both of the above

- Neither of the above

In the word classification task, an independent sample of participants was recruited (N = 80). Participants were presented with the 100 words with the highest weight in the

definitions of indebtedness collected in the large-scale sampling and were instructed as following:

> *In this experiment, you will see a list of words that people use in everyday helping and receiving helping situations. After seeing each word, please classify it into one of the following five categories according to your understanding:*
>
> *1. Appraisal: if you think that the word is related to appraisals or evaluations on the benefits, costs, intentions or other factors in the situation, please classify the word into this category.*
>
> *2. Emotion (Feeling): if you think that this word is related to or reflects certain emotions and feelings of an individual in the situation, please classify the word into this category.*
>
> *3. Behavior: if you think the word is related to the behavior of an individual in the situation, please classify the word into this category.*
>
> *4. Person: if you think this word is related to person, please classify the word into this category.*
>
> *5. Other: if you think the word has nothing to do with any of the four categories above, please classify the word into this category.*

Study 2 (behavioral study)

In Study 2a and Study 2b, seven participants came to the experiment room together. An intra-epidermal needle electrode was attached to the left wrist of each participant for cutaneous electrical stimulation [5]. The first pain stimulation was set as 8 repeated pulses, each of which was 0.2 mA and lasted for 0.5 ms. A 10-ms interval was inserted between pulses. Then we gradually increased the intensity of each single pulse until the participant reported 6 on an 8-point pain scale (1 = not painful, 8 = intolerable). The participants reported that they could only experience the whole train of pulse as a single stimulation, rather than as separate shocks. The final intensity of pain stimulation was calibrated to a subjective pain rating of "6", which was a moderate punishment for the participants.

Both Study 2a and Study 2b consisted of 3 sessions. All stimuli were presented using

PsychToolBox 3.0.14 (www.psychtoolbox.org) in Matlab 2016a (Mathworks, Natick, MA, USA). Participants were instructed as following:

In this experiment, you will play an interpersonal game, which is composed of two roles: the Decider (Player A) and the Receiver (Player B). The Receiver will be in some trouble and the Decider can decide whether to help the Receiver at the cost of his/her own interests. Several previous participants have come to our lab during Stage 1 of our study and made decisions as the Deciders. Now this experiment belongs to Stage 2 of this study, which consists of three sessions. In the first two sessions of the experiment, you will perform as the Receiver, facing the decisions made by each previous Decider in Stage 1 and make your own decisions. Then in the last session, you will perform as the Decider and make decisions about whether to help the Receivers in the next stage.

During Session 1 (the main task), each participant played multiple single-shot rounds of this interpersonal game as a Receiver (Fig. 1C). In each round, the participant was to receive a 20s pain stimulation with the intensity of 6, and was paired with a same-sex Decider in Stage 1. The participant was instructed that the Decider in each trial was distinct from the ones in any other trials and only interacted with the participant once during the experiment. Each Decider was informed of the participant's situation and was endowed with 20 yuan (~ 3.1 USD). The Decider could decide whether he/she want to spend some of the endowment to reduce the duration of the participant's pain. The more the Decider spent, the more the duration of the participant's pain reduced. The maximum pain reduction is 16 seconds. In fact, each Decider's decision was pre-determined by the computer program (Table S2).

Each trial began by informing the participant that the program had randomly chosen one of the participants in Stage 1 as the Decider for the current trial, with the experiment number and a blurred picture of this Decider presented in the screen. Then the Decider's decision on how much he/she decided to spend to help the participant was presented. After presenting the Decider's decision, the participant indicated how

6

much he/she thought this Decider expected him/her to reciprocate (i.e., second-order belief of the co-player's expectation for repayment; continuous rating scale from 0 to 25 using mouse, step of 0.1 yuan). In half of the trials, the participant had to accept the Decider's help; in the remaining trials, the participant could decide whether to accept the Decider's help or not. If the participant accepted the help, the Decider's cost and the participant's pain reduction in this trial would be realized as the Decider's decision; if the participant did not accept the help, the Decider's would spent no money and the duration of participant's pain stimulation would be the initial 20s. At the end of each trial, the participant was endowed with 25 yuan (~ 3.8 USD) and decided how much he/she want to allocate to the Decider as reciprocity in this trial from this endowment (continuous choice from 0 to 25 using mouse, step of 0.1 yuan).

We manipulated the intention of the Decider (i.e., the benefactor) by providing participants with extra information regarding the Decider's expectations of reciprocation (i.e., **extra information about benefactor's intention**) below Decider's number at the beginning of each trial. Each participant were instructed that before making decisions, some Deciders were informed that the participant would be endowed with 25 yuan and could decide whether to allocate some endowments to them as reciprocity (i.e., *Strategic condition*), whereas the other Deciders were informed that the participant had no chance to reciprocate after receiving help (i.e., *Altruistic condition*). The endowment that each Decider spent to help the participant (i.e., **Benefactor's cost**) was parametrically manipulated in each condition, which enabled us to predict participants' behaviors using computational modeling.

In Study 2b, to dissociate the effect of Benefactor's cost and participant's benefit (i.e., pain reduction), we further manipulated the exchange rate between the Decider's cost and participant's pain reduction (i.e., **Efficiency**, 0.5, 1, and 1.5) on the basis of Study 2a where Efficiency was 1. Thus, the participant's pain reduction was calculated by: Pain reduction = Decider's cost / Decider's endowment × Efficiency × Maximum

pain reduction (16s). For both Study 2a and Study 2b, each condition included one trial for each Benefactor's cost – Efficiency combination (see Table S2 for detailed information about the experimental design).

At the end of the experiment, 5 trials in Session 1 were randomly selected to be realized. The participant received the average pain stimulation in these 5 trials. The participant's final payoff was the average amount of endowment the participant left for him/herself across the chosen trials. The final payoff of each Decider was the average amount of endowment the Decider left plus the amount of endowment the participant allocated to him/her. Participants were informed of this arrangement before the execution of the experiment.

During Session 2, all of the decisions in the first session were displayed again in a random order. After being shown the Decider's information and his/her decision, the participant was asked to recall their feelings when they received the help of the Decider and to make the following ratings. To rule out the concern that the ratings of gratitude and indebtedness influence each other, the orders of ratings of gratitude and indebtedness were counter-balanced across trials.

- "How much gratitude do you feel for this Decider's decision?" (Gratitude)
- "How much indebtedness do you feel for this Decider's decision?" (Indebtedness)
- "How much do you think this decider care about you?" (Perceived care)
- "How much afraid/pressure do you feel for the decider's expectation for repayment?" (Sense of obligation)
- "How much guilt do you feel for this Decider's decision?"(Guilt)

At Session 3, each participant took a selfie and was informed that this photo would be blurred and used in our future experiments. Then they made 6 decisions as a Decider without any information about whether the future Receiver could reciprocate or not. They were told that if their decisions in this session were selected for future

8

experiments, they would be paid according to their own decisions and future Receivers' decisions by Alipay (an online payment method in China). This procedure was to ensure that participants believed in experimental design.

Study 3 (fMRI study)

Each participant came to the scanning room individually. The pain-rating procedure and the three sessions of the task in the fMRI study were the same as Study 2a, except that in the main task of fMRI study (Session 1), the trials in which participants could decide whether to accept help were excluded. Therefore, participants had to accept the Decider's help throughout the experiment. Session 1 (the main task) was conducted in the fMRI scanner, while Sessions 2 and 3 were conducted outside the scanner after MRI scanning.

For Session 1, the manipulation on extra information regarding the benefactor's intention (Strategic condition and Altruistic condition) in the fMRI study was the same as Study 2a. The endowment that the Decider spent to help the participant (i.e., Benefactor's cost) in each condition was parametrically manipulated which enabled us to predict participants' behaviors using computational modeling. Specifically, the participant was informed that each Decider could only spend even amounts of the endowment to help the participant. Therefore, the final Benefactor's cost had 9 levels (4, 6, 8, 10, 12, 14, 16, 18, and 20, see Table S2 for details). The scanning session consisted of three runs and lasted for about 39 min. Each run lasted for 13 min and consisted of 18 trials (including the 9 levels of the benefactor's cost in Strategic condition and Altruistic conditions respectively), the order of which were pseudorandomized.

Each trial began by informing the participant that the program had randomly chosen one of the participants in the Stage 1 as the Decider for the current trial with the experiment number and the blurred picture of this Decider, and information of whether this Decider knew that the participant could or could not repay was presented

on the screen (Information period, lasted for 4 seconds). Then the Decider's decision on how much to spend to help the participant was presented (Outcome period, lasted for 5 seconds). After presenting the Decider's decision, Participants indicated how much he/she thought this Decider expected him/her to reciprocate (i.e., second-order belief of the co-player's expectation for repayment; rating scale from 0 to 25 using left and right buttons to move the cursor, step of 1 yuan). Participants needed to make choice within 8 seconds, and then their choices continued to be presented until the total time reached 8 seconds ($Belief_{2nd}$ rating period). After that the participant had to accept the Decider's help, was endowed with 25 yuan (~ 3.8 USD) and decided how much to allocate to the Decider as reciprocity in this trial (from 0 to 25 using left and right buttons to move the cursor, step of 1 yuan). Participants needed to make the choice within 8 seconds, and then their choices continued to be presented until the total time reached 8 seconds (Allocation period). Before and after each period, a fixation cross was presented for a variable interval ranging from 2 to 6 s, which was for the purpose of fMRI signal deconvolution.

Additional measures

Each participant was asked to complete Communal Orientation Scale and Exchange Orientation Scale [6,7] after the experiments in Studies 2 and 3. These two separate and orthogonal scales measure individual differences in tendencies to follow communal and exchange norms in everyday life. In the current study, the individuals' relative communal and relative exchange orientation was defined as

$$\frac{Score\ on\ \text{Communal Orientation Scale}}{Sum\ of\ scores\ on\ \text{Communal and Exchange Orientation Scales}} \text{ and}$$

$$\frac{Score\ on\ \text{Exchange Orientation Scale}}{Sum\ of\ scores\ on\ \text{Communal and Exchange Orientation Scales}}, \text{ respectively.}$$

Behavioral data analyses

Behavioral data analyses were carried out in RStudio Version 1.1.383 [8] and IPython/Jupyter Notebook (Python 3.6.8) [9], and was plotted using seaborn 0.9.0 (https://seaborn.pydata.org/index.html). All of the three studies employed Akaike

information criterion (AIC) [10] for model comparison. AIC information criterion reflects both goodness of fit and complexity of the model. The smaller the AIC, the better the model.

Study 1

The definition of indebtedness. We conducted three different analyses to clarify the definition of indebtedness. First, using the emotional ratings for daily-life events of receiving and rejecting help, we tested: (1) whether both guilt and obligation ratings independently explained indebtedness ratings, and (2) whether the model with both guilt and obligation ratings as predictor for indebtedness ratings were better than models with single predictors. We conducted between-participant linear regressions with indebtedness ratings as an independent variable. Three models were included and compared. Model 1 included obligation rating as the single predictor. Model 2 included guilt rating as the single predictor. Model 3 included both obligation and guilt ratings as predictors. To rule out the possibility that these emotions might covary with other variables (e.g., benefactor's cost, the participant's benefit and the social distance between the participant and the benefactor), we re-conducted model comparison after controlling for the effects of self-reported benefactor's cost, the participant's benefit and the social distance between the participant and the benefactor. Results remained the same after controlling for these variables (Table S1).

Second, we calculated the frequency of choosing each of the four options in the question "In daily life, what do you think is/ are the source(s) of indebtedness?", as well as the frequencies that participants thought "Negative feeling for harming the benefactor" and "Negative feeling for the pressure to repay caused by other's ulterior intentions" as the sources of indebtedness (i.e., the frequency of choosing each single option plus the frequency of choosing "Both of the above").

Third, for the self-reported definition of indebtedness, we used IPython/Jupyter Notebook [9] in combination with "Wordcloud"

(https://amueller.github.io/word_cloud/index.html) and "Jieba" (https://github.com/fxsjy/jieba) packages to conduct text segmentation. The weight of each word in the whole text was calculated using Term Frequency-Inverse Document Frequency (TF-IDF) [11,12]. This method calculates the importance of a word in the whole corpus based on the frequency of its occurrence in the text and the frequency of its occurrence in the whole corpus. The advantage of this method is that it can filter out some common but irrelevant words, while retaining important words that affect the whole text. Using this method, the 100 words with the highest weight/frequency in the definitions of indebtedness were extracted (Appendices S2). To extract the words related to emotions and feelings in the definition of indebtedness, these 100 words were classified by an independent sample of participants (N = 80) into levels of appraisal, emotion, behavior, person and other. Then we conducted Latent Dirichlet Allocation (LDA) based topic modeling on the emotional words of indebtedness using collapsed Gibbs sampling implemented in "lda" package (https://lda.readthedocs.io/en/latest/) [13]. LDA is a generative probabilistic model for collections of discrete data such as text corpora, which is widely used to discover the topics that are present in a corpus [14]. In LDA, documents are represented as random mixtures over latent topics, and each topic is then characterized by a distribution over words. The goal of LDA is to map all the documents to the topics in a way, such that the words in each document are mostly captured by those imaginary topics. Because Chinese retains its own characters of various structures, the pre-processing including the exclusion of stop words and synonym combination are implemented before topic modeling[15]. We conducted model comparison on topic models with topic numbers ranged from 2 to 10 to test whether the two-topic solution performed the best. We then extracted the top two topics from the emotional words in the definition of indebtedness to test whether the words in these two topics fits our hypothesis that participants' definition of indebtedness contain both the words related to guilt and obligation.

The influences of emotions on behaviors. First, we combined the data of the events of

receiving help (Part 1) and rejecting help (Part 2) in the questionnaire, took whether the participant rejected help in the event (reject was coded as 1, and accept was coded as 0) as the dependent variable, and used logistic regression to examine the influences of gratitude, indebtedness, guilt, the sense of obligation on decisions of whether to accept help respectively. Second, we used data of Part 1 in the questionnaire in which participants reported events of receiving help and made ratings on the need to reciprocate as the dependent variable, and used linear regression to examine the influences of gratitude, indebtedness, guilt, the sense of obligation on reciprocity after receiving help.

Study 2 and Study 3

The definition of indebtedness. In studies 2 and 3, using post-experiment trial-by-trial emotional ratings for each participant, we fit general linear mixed models (LMM) for indebtedness rating including by-participant and by-experiment random slopes for each fixed effect. LMMs were conducted using lme4 package [16] in R. For all studies, three models were included and compared. Model 1 included obligation rating as the single predictor. Model 2 included guilt rating as the single predictor. Model 3 included both obligation and guilt ratings as fixed effects. Given that these three emotions might be modulated by experimental variables (e.g., benefactor's cost, extra information about the benefactor's intention and efficiency), which might result in the covariance of these variables and hide the true relationships between them. To exclude this possibility, the effects of experimental variables were regressed out from obligation, guilt, and indebtedness ratings by conducting three LMMs with each of these three variables as dependent variable separately and the experimental variables as fixed effects. By-participant and by-experiment random slopes for each fixed effect were included in each LMM. Residuals of each LMM were then extracted as the index for each emotion after controlling for the effects of experimental variables and were used to re-conduct model comparison. Results remained the same after controlling for the experimental variables (Table S1).

The effect of the extra information about benefactor's intention, benefactor's cost, and efficiency on participants' appraisal, emotional and behavioral responses. To test the effects of the benefactor's cost and extra information about benefactor's intention on beneficiary's appraisals (second-order belief and perceived care), emotions (gratitude, indebtedness, guilt and obligation) and behaviors (reciprocity and whether reject help), in Study 2a we conducted linear mixed model (LMM) analyses for each dependent variable separately with participant as a random effect. By-subject random slopes for each fixed effect were also included in the models [17]. For each dependent variable, in addition to the full model that contained the benefactor's cost, extra information about benefactor's intention, and the interaction terms as predictors, we tested different reduced models with different combinations of independent variables and interaction terms as predictors and conducted model comparisons to test whether any reduced model better fitted participants' responses. The effects of predictors in the winning model were finally reported (Table S3-1). Similar model reduction procedures were conducted for Study 2b (Table S3-2), the combining data of Studies 2a and 2b (Table S3-3) and Study 3 (Table S4).

Relationships between appraisals and emotions. Given the correlations between appraisals and emotions (Fig. 3D), we conducted principal component analysis (PCA) to reduce the dimension of data and examine the relationship between appraisals and emotions. All the variables were centered within participant to exclude the influences of individual differences in the range of ratings. PCA was performed in an iterative process using the "nFactors" package [18] and "psych" package [19] for R. To determine the number of components to retain, the correlation matrix between the 6 variables was submitted to a parallel analysis [20]. Parallel analysis performs a principal factor decomposition of the data matrix and compares it to a principal factor decomposition of a randomized data matrix. This analysis yields components whose eigenvalues (magnitudes) are greater in the observed data relative to the randomized data. The nScree function was used to determine the number of components to retain with four retention criteria: the Kaiser-Guttman, the scree test (optimal coordinates and

acceleration factor), and the parallel analysis. Finally, the function plotnScree was used to plot the four factor retention criteria. Three criteria pointed to a two-component solution except for the acceleration factor (Fig. S2E). Thus, we extracted principle components (PCs) using the function Principal by restricting the number of PCs to two.

Given that all the appraisals and emotions were modulated by experimental variables (i.e., extra information about benefactor's intention, benefactor's cost, and help efficiency), which might result in the covariance of these variables and hide the true relationships between them. To exclude this possibility, the effects of experimental variables were regressed out from second-order belief, perceived care, obligation, guilt, indebtedness and gratitude ratings by conducting 6 LMMs with each of these six variables as dependent variable separately and Benefactor's cost, Efficiency and intention condition as fixed effects. By-participant and by-experiment random slopes for each fixed effect were included in each LMM. Residuals of each LMM were then extracted as the index for each appraisal or emotion after controlling for the effects of experimental variables and were used to re-conduct PCA analysis. Results remained the same after controlling for the experimental variables.

Computational Modeling

Based on the theoretical model, we applied a computational modeling approach to formally test the hypotheses derived from our theoretical model and quantitively capture the weights on communal motivation and obligation motivation behind behaviors. Integrating previous models of other-regarding preference [21-23], our computational model first posits that, when a participant face a strategy/decision of the benefactor (player A), his/her (player B's) utility (U) of a strategy/decision (D_B, i.e., the amount of reciprocity and accepting or rejecting help) in response to the help results from a trade-off between financial self-interest (π) and social motivations, weighted by a greed parameter Θ ($0 < \Theta < 1$). As illustrated in the conceptual model, the social motivations include the communal motivation and the obligation motivation,

and the parameter Φ captures individual weight on communal motivation and $1 - |\Phi|$ captures individual weight on obligation motivation. The participant (player B) thus makes decision by maximizing the utility of this decision in the following utility function (Eq. S1):

$$U(D_B) = \theta_B * \pi_B + (1 - \theta_B) * (\phi_B * U_{Communal} + (1 - |\phi_B|) * U_{Obligation})$$ **Eq. S1**

Since both feelings of communal motivation and obligation motivation are likely to motivate player B to reciprocate after receiving favors [24-29], for the model of reciprocity, Φ is defined as a free parameter between 0 and 1, which captures the trade-off between communal motivation and obligation motivation in reciprocity. The more Φ approximates to 1, the more the participant cares about the communal motivation during reciprocity; the more Φ approaches 0, the more the participant cares about the obligation motivation during reciprocity.

In contrast to reciprocity, the decisions of whether to accept or reject help might be more complex. The sense of obligation may motivate rejecting the help to avoid being in the benefactor's debt [25,26,30]. For the communal motivation, while gratitude may motivate one to accept help to build a communal relationship [31,32], guilt may motivate one to reject help to avoid burdening a benefactor [33,34]. We capture this complexity by extending the range of Φ to [-1, 1]. Specifically, $\Phi < 0$ indicates that the communal motivation motives the participant to reject the help, while $\Phi > 0$ indicates that the communal motivation motives the participant to accept the help. Correspondingly, individual weight on obligation motivation is captured by $1 - |\Phi|$, which ranges from 0 to 1, indicating that the obligation motivation motivates individual to reject help. This model performed better than the model restricted the range of Φ to [0, 1].

Modeling of each utility term. We first modeled the utility of self-interest (π_B) as Eq. S2. For each amount of reciprocity (D_B), the self-interest was defined as the percentage of money the participant obtained from the total endowment (γ_B). For the decisions of whether to accept, the self-interest from the decision of accepting help was defined as the as the percentage of pain reduction from the total amount of the maximum pain reduction, which depended on the how much the benefactor spent to

help (D_A) and the exchange rate between the benefactor's cost and the participant's benefit (μ).

$$\pi_B = \begin{cases} \frac{\gamma_B - D_B}{\gamma_B} & \text{Reciprocity} \\ \frac{D_A * \mu}{\max(D_A * \mu)} & \text{Accept/Reject help} \end{cases}$$

Eq. S2

Here, instead of modeling each specific emotion, we aimed to test whether and how individuals make trade-off between communal motivation (i.e., guilt and gratitude) and obligation motivation (i.e., sense of obligation) to make decisions regarding receiving favors. Therefore, we modeled and used the perceived care (ω_B) and the second-order belief of the benefactor's expectation for repayment (E_B'') as the indexes for communal and obligation motivations according to the results of participants' ratings (Fig.3, B-C), since we hypothesized theses two appraisals were the key factors that induced dual motivations. The participant's second-order belief of how much the benefactor expected in each trial was defined as a function of the extra information about benefactor's intention and benefactor's cost (D_A) (Eq. S3). This indicted that, when the benefactor knew that the participant could not reciprocate after receiving help, the participant would think that the benefactor expected no repayment after the help. However, when the benefactor knew that the participant could reciprocate after receiving help, the higher the benefactor's cost, the more repayment the participant's thought the benefactor would expect.

$$E_B'' = \begin{cases} 0 & \text{Altruistic condition} \\ D_A & \text{Strategic condition} \end{cases}$$

Eq. S3

The participant's perceived care (ω_B) in each trial was defined as a function of the benefactor's cost and second-order belief (Eq. S4). Specifically, we assumed that the perceived care from the help increased as a linear function of how much the benefactor spent (D_A) from his/her endowment (γ_A); however, this effect was reduced by the second-order belief of the benefactor's expectation for repayment (E_B''). Here, the parameter kappa (κ) is a free parameter ranges from 0 and 1, representing to what extent the second-order belief of benefactor's expectation for repayment would reduce the participant's perceived care.

$$\omega_B = \frac{D_A - \kappa_B * E_B''}{\gamma_A}$$

Eq. S4

To avoid the potential correlation between these two items, which would impact the quality of modeling fitting, we took the amount of reciprocity as the dependent variable and tested the degree of multicollinearity between ω_B and E_B'' using variance inflation factor (VIF). The VIF reflects the degree to which the variance of the regression coefficient in multiple regressions is magnified due to multicollinearity. VIF equaling 5 is commonly regarded the threshold to assess the level of multicollinearity [35]: a VIF below 5 indicates the degree of multicollinearity is acceptable, while a VIF more than 5 indicates the degree of multicollinearity should be carefully considered. To be noted, when testing the relationship between perceived care and second-order belief, we controlled for the effects of experimental variables to investigate the true relationship between these two variables. On the contrary, in the computational modeling capturing how individuals made trade-off between these two motivations, the variability of experimental variables could help to reduce the inherent negative correlation between ω_B and E_B'' and benefited model fitting. Results demonstrated a low level of multicollinearity between ω_B and E_B'' (Study 2a: VIF = 1.06; Study 2b: VIF = 2.30).

Then we applied the two appraisals as the indexes of communal motivation and obligation motivation, and defined $U_{Communal}$ and $U_{Obligation}$ as the functions of ω_B and E_B'' respectively, following the hypotheses for the decision-making process of whether to accept help and reciprocity after receiving help (Eq. S5 and Eq. S6).

$$U_{Communal} = \begin{cases} -(\frac{\omega_B * \gamma_B - D_B}{\gamma_B})^2 & \text{Reciprocity} \\ \omega_B & \text{Accept/Reject help} \end{cases}$$ **Eq. S5**

$$U_{Obligation} = \begin{cases} -(\frac{E_B'' - D_B}{\gamma_B})^2 & \text{Reciprocity} \\ -\frac{E_B''}{\gamma_B} & \text{Accept/Reject help} \end{cases}$$ **Eq. S6**

Specifically, for reciprocity, $U_{Communal}$ and $U_{Obligation}$ were defined as the quadratic functions of ω_B and E_B''. The function of $U_{Communal}$ assumed that the participant maximized the utility of communal by minimizing the difference between the amount of reciprocity (D_B) and the amount of money the participant willing to give to

reciprocate the benefactor's kindness, which depended on the perceived care (ω_B) and the amount of endowment the participant had (γ_B). The function of $U_{Obligation}$ assumed that the participant maximized the utility of obligation by minimizing the difference the amount of reciprocity (D_B) and the second-order belief of how much the benefactor expected (E_B''). For decisions of whether to reject help, $U_{Communal}$ and $U_{Obligation}$ were defined as the linear functions of ω_B and E_B'', which indicated that the higher the perceived care, the higher utility of accepting the help, whereas the higher the second-order belief of the benefactor's expectation for repay, the higher utility of rejecting the help.

Therefore, the final utility functions for the amount of reciprocity after receiving help and the decisions of whether to accept help were defined as Model 1.1 and Model 2.1 respectively:

Model 1.1:

$$U(D_B) = \theta_B * \frac{\gamma_B - D_B}{\gamma_B} - (1 - \theta_B) * (\phi_B * (\frac{\omega_B * \gamma_B - D_B}{\gamma_B})^2 + (1 - \phi_B) * (\frac{E_B'' - D_B}{\gamma_B})^2)$$

Model 2.1:

$$U(Accept) - U(Reject) = \theta_B * \frac{D_A * \mu}{\max(D_A * \mu)} + (1 - \theta_B) * (\phi_B * \omega_B - (1 - |\phi_B|) * \frac{E_B''}{\gamma_B})$$

See the following table for the meaning of each symbol of variables.

Variable	Meaning
D	Decision over choice space, e.g., the benefactor's cost, the beneficiary's amount of reciprocity and the beneficiary's decision of accepting help
θ	Greed sensitivity
ϕ	Mixture weight of $U_{Communal}$ and $U_{Obligation}$
γ	Endowment size
π	Self interest
μ	The efficiency of help

Variable	Meaning
E''	Second-order belief of how much the benefactor expects
ω	Perceived care
κ	The influence of second-order belief (E'') on perceived care (ω) (higher indicates lower perceived care)
λ	Inverse temperature parameter
n	Number of trials
t	Trial

Model fitting. Considering that the amount of reciprocity is a continuous variable, and the decision of whether to accept help is a binary variable, we applied different link functions relating utility functions to behaviors for these two lines of decisions.

For reciprocity, we estimated the model parameters by minimizing the sum of squared error between the model's behavioral prediction and actual behavior over all the trials that participants had to accept help using Matlab's fmincon routine. We minimized the following objective function for each participant,

$$SSE = \sum_{t=1}^{n}(\frac{D_B(t) - max(U(D_B(t)))}{\gamma_B} * 100)^2$$

Eq. S7

with *t* indicating trial number. To avoid ending the fitting procedure at a local minimum, the model-fitting algorithm was initialized at 1000 random points in theta-phi-kappa parameter space for each participant.

For decisions of whether to accept or reject help, the model was calibrated using a softmax specification with inverse temperature parameter λ using maximum likelihood, such that in each trial, the probability of the participant choosing one option is given by

$$P(D_B) = \frac{1}{1 + e^{-\lambda * U(D_B)}}$$

Eq. S8

We then conducted maximum likelihood estimation at the individual level by minimizing the negative log likelihood function over each trial *t* with 1000 different

starting values:

$$LLE = -\sum_{t=1}^{n} log(P(D_B(t)))$$

Eq. S9

Model comparison for reciprocity. First, we compared the quadratic version of model (Model 1.1) with 2 models with linear versions of $U_{Communal}$ and $U_{Obligation}$ (Model 1.2 to Model 1.3). These linear models had mathematical representations different from the quadratic model but shared the same hypotheses of decisions-making, such that the participant maximized the utility of communal motivation by minimizing the difference between the amount of reciprocity and the amount of money the participant were willing to give to reciprocate the benefactor's kindness, and maximized the utility of obligation motivation by minimizing the difference between the amount of reciprocity and the second-order belief of how much the benefactor expects.

Model 1.2:

$$U(D_B) = \theta_B * (\gamma_B - D_B) - (1 - \theta_B) * (\phi_B * \max(\omega_B * \gamma_B - D_B, 0) + (1 - \phi_B) * \max(E_B'' - D_B, 0))$$

Model 1.3:

$$U(D_B) = \theta_B * (\gamma_B - D_B) - (1 - \theta_B) * (\phi_B * |\omega_B * \gamma_B - D_B, 0| + (1 - \phi_B) * |E_B'' - D_B, 0|)$$

To examine the necessities of both communal motivation and obligation motivation, we compared the full model (Model 1.1) with the model without obligation item (Model 1.4) and without communal motivation item (Model 1.5).

Model 1.4:

$$U(D_B) = \theta_B * \frac{\gamma_B - D_B}{\gamma_B} - (1 - \theta_B) * (\frac{\omega_B * \gamma_B - D_B}{\gamma_B})^2$$

Model 1.5:

$$U(D_B) = \theta_B * \frac{\gamma_B - D_B}{\gamma_B} - (1 - \theta_B) * (\frac{E_B'' - D_B}{\gamma_B})^2$$

We further compared our model assuming participants made trade-off between

21

communal and obligation motivations (Model 1.1) with two models that had with separate parameters for different terms (Model 1.6 and Model 1.7), in which the ranges of parameters were defined as [0, 1].

Model 1.6:

$$U(D_B) = \theta_B * \frac{\gamma_B - D_B}{\gamma_B} - (1 - \theta_B) * (\phi_{B1} * (\frac{\omega_B * \gamma_B - D_B}{\gamma_B})^2 + \phi_{B2} * (\frac{E_B'' - D_B}{\gamma_B})^2)$$

Model 1.7:

$$U(D_B) = \theta_B * \frac{\gamma_B - D_B}{\gamma_B} - \phi_{B1} * (\frac{\omega_B * \gamma_B - D_B}{\gamma_B})^2 - \phi_{B2} * (\frac{E_B'' - D_B}{\gamma_B})^2$$

Another possibility was that participants did not make decisions based on dual motivations, but decided how much to reciprocate simply according to the benefactor's cost. The closer the amount of reciprocity to the benefactor's cost, the larger the utility of this action. In Altruistic and Strategic conditions, the weights on the benefactor's cost in the total utility were different, capturing by parameters Φ_1 and Φ_2 respectively, resulting in the behavioral differences under different conditions (Model 1.8). In this model, C is the condition-indicating coefficient: $C = 1$ represents Altruistic condition, and $C = 0$ represents the Strategic condition.

Model 1.8:

$$U(D_B) = \theta_B * \pi_B + (1 - \theta_B) * (C * \phi_{B1} * U_{Cost} + (1 - C) * \phi_{B2} * U_{Cost})$$

$$U_{Cost} = -(\frac{D_A - D_B}{\gamma_B})^2$$

The third possibility is that participants made decisions according to inequity aversion model [22,36], which assumed that participants cared not only self-interest, but also the payoff difference between self and other (Model 1.9). The smaller the difference between self-payoff and the benefactor's payoff (or the closer the reciprocity is to 1/2 of their total payoff), the greater the utility of this action. In Altruistic and Strategic conditions, the weights on inequity were different, captured by parameters Φ_1 and Φ_2 respectively, resulting in the behavioral differences under different conditions.

Model 1.9*:*

$$U(D_B) = \theta_B * \pi_B + (1 - \theta_B) * (C * \phi_{B1} * U_{Inequity} + (1 - C) * \phi_{B2} * U_{Inequity})$$

$$U_{Inequity} = -(\frac{\gamma_B - D_B}{\gamma_A - D_A + \gamma_B} - \frac{1}{2})^2$$

Model comparison for decisions of whether to accept or reject the help. Similar to reciprocity, to examine the necessities of both communal and obligation motivations, we compared the full model (Model 2.1) with the model without obligation item (Model 2.2) and the one without communal motivation item (Model 2.3).

Model 2.2:

$$U(Accept) - U(Reject) = \theta_B * \frac{D_A * \mu}{\max(D_A * \mu)} + (1 - \theta_B) * \omega_B$$

Model 2.3:

$$U(Accept) - U(Reject) = \theta_B * \frac{D_A * \mu}{\max(D_A * \mu)} - (1 - \theta_B) * \frac{E_B^{''}}{\gamma_B}$$

The softmax function for binary choice of whether to accept help added one more temperature parameter λ to estimate, which may reduce the quality of model fitting due to the limited the number of trials. Therefore, we introduced individual κ obtained from the model for reciprocity (Model 1.1) to reduce the number of parameters, and re-estimated Model 2.1 and Model 2.2 (labeled as Model 2.4 and Model 2.5). Model 2.3 with no κ in the model was not re-estimated. We further compared the outperformed model among all models above (Model 2.4), which assumed that participants made trade-off between communal and obligation motivations (Model 2.4) with two models that had separate parameters for different terms (Model 2.6 and Model 2.7), in which the ranges of Θ and Φ_2 were defined as [0,1], and the ranges of Φ_1 was defined as [-1,1].

Model 2.6*:*

$$U(Accept) - U(Reject) = \theta_B * \frac{D_A * \mu}{\max(D_A * \mu)} + (1 - \theta_B) * (\phi_{B1} * \omega_B - \phi_{B2} * \frac{E_B^{''}}{\gamma_B})$$

Model 2.7:

$$U(Accept) - U(Reject) = \theta_B * \frac{D_A * \mu}{\max(D_A * \mu)} + \phi_{B1} * \omega_B - \phi_{B2} * \frac{E_B''}{\gamma_B}$$

To test the necessities of extending the range of *Φ to* [-1, 1], we reduced the range of Φ *to* [0, 1] (Model 2.8) and tested whether the goodness of model fitting decreased.

Because in the decision of whether to accept help, as long as the participant accepted help, the cost of the benefactor would increase, and the degree of inequity between the participant and the benefactor would also increase. Therefore, the function of " Reciprocity according to benefactor's cost model" was the same as that of the inequity aversion model. In inequity aversion model (Model 2.9), in Altruistic and Strategic conditions, the weights on inequity were different, captured by parameters $Φ_1$ and $Φ_2$ respectively, resulting in the behavioral differences under different conditions. *C* is the condition-indicating coefficient: *C* = 1 represents Altruistic condition, and *C* = 0 represents the Strategic condition.

Model 2.9:

$$U(Accept) - U(Reject) = \theta_B * \pi_B + (1 - \theta_B) * (C * \phi_{B1} * U_{Inequity} - (1 - C) * \phi_{B2} * U_{Inequity})$$
$$U_{Inequity} = \frac{D_A}{\gamma_A}$$

Parameter recovery. We ran parameter recovery analyses to ensure that our model was robustly identifiable [37]. To this end, we simulated data for all models for each participant using the original trials settings, model formulations and original parameters estimated from the behavioral data. The number of simulations is the same as the trial number used to fit the model. We refitted the model using 1000 random start locations to minimize the possibility of the algorithm getting stuck in a local minimum. We then assessed the degree to which the parameters could be recovered by calculating the similarity between all the parameters estimated from the behavioral data and all the parameters estimated from the simulated data using a Pearson correlation.

Model predictions. To test whether our model predicted participants' reciprocity behaviors, LMM was conducted with actual reciprocity as dependent variable including by-participant and by-experiment random slopes for the fixed effect of model prediction. R^2 of this LMM was extracted using sem.model.fits function implemented in the piecewiseSEM package in R.

FMRI Data acquisition and preprocessing

Images were acquired using a 3-T Prisma Siemens scanner (Siemens AG, Erlangen, Germany) with a standard head coil at Peking University (Beijing, China). T2-weighted echoplanar images (EPI) were obtained with blood oxygenation level-dependent (BOLD) contrast. Sixty-two transverse slices of 2.3 mm thickness that covered the whole brain were acquired using multiband EPI sequence in an interleaved order (repetition time = 2000 ms, echo time = 30 ms, field of view = 224×224 mm^2, flip angle = 90°). The fMRI data were preprocessed and analyzed using Statistical Parametric Mapping software SPM12 (Wellcome Trust Department of Cognitive Neurology, London). Images were slice-time corrected, motion corrected, resampled to 3 mm × 3 mm × 3 mm isotropic voxels, normalized to MNI space, spatially smoothed with an 8 mm FWHM Gaussian filter, and temporally filtered using a high-pass filter with a cutoff frequency of 1/128 Hz.

FMRI data analysis

Whole-brain analysis. We were specifically interested in brain processes during the Outcome period, in which participants learned about the benefactor's decision to help. Using a model-based fMRI analytic approach [38], we fit three separate general linear models (GLMs 1-3) to each voxel's timeseries to identify brain regions that tracked different components of the computational model. These included values of: (1) reciprocity (GLM 1), (2) communal motivation, which depended on the perceived care from the help (ω_B) (GLM 2), and (3) obligation motivation, which depended on the second-order belief of the benefactor's expectation for repayment (E_B'') (GLM 3).

As our model hypothesizes that communal and obligation motivations arise from the perceived care from the help (ω_B) and the second-order belief of the benefactor's expectation for repayment (E_B'') respectively, we used ω_B and E_B'' in the computational model as indices for communal and obligation motivations and conducted parametric analyses. Brain responses to ω_B and E_B'' reflected how much information in neural patterns was associated with each motivation in the brain. An alternative approach is to use the $U_{Communal}$ and the $U_{Obligation}$ from our computation model as parametric modulators when estimating brain responses. However, in our model, $U_{Communal}$ and the $U_{Obligation}$ were defined as negative quadratic functions, the maximum values of which were zero. As we predicted and observed, participants behaved to maximize their $U_{Obligation}$ by minimizing the differences between the amount of reciprocity and E_B'', and to maximize their $U_{Communal}$ by minimizing the differences between the amount of reciprocity and ω_B. Therefore, in a large proportion of trials, the $U_{Obligation}$ and $U_{Communal}$ were near zero as a result of participant's decisions, making them inefficient for parametric analysis to capture how successfully participants behaved in accordance with their motivations. In contrast, ω_B and E_B'' better captured the inferences that comprised participants' motivations and were more suitable for testing our hypotheses about brain responses.

In GLM 1 for the processing of reciprocity, we modeled the Outcome period as the key regressor, starting from the time the benefactor's choice was revealed and spanning the duration of this event, which was 5s. The amount of reciprocity in each trial was added as a parametric modulator on this regressor. Regressors of no interest included Information period (onset of the presentation of the benefactor's picture and extra information regarding intention, 4s), Second-order belief rating period (starting from the time the rating screen presented and spanning to the time that the participant made choice), Allocation period (starting from the time the rating screen presented and spanning to the time that the participant made choice), and Missed response (the missing decision period for Belief$_{2nd}$ or allocation, 8s). Six movement parameters were included as regressors of no interest. In this GLM, one contrast was defined as

the positive effect of the parametric modulator. GLM 2 for the processing of communal motivation was the same as GLM 1 except that model-driven communal motivation (ω_B) in each trial was added as parametric modulators on the key regressor.

As our model predicted that the participants' sense of obligation was near zero in Altruistic condition and increased linearly in Strategic condition with the increase of the benefactor's cost; this pattern was not a normally distribution. Therefore, in GLM 3, instead of linear parametric analysis, to identify brain regions involved in obligation processing, we modeled the Outcome period as four separate regressors: Strategic_Lowcost (including trials where benefactor's cost = 4, 6, or 8 in Strategic condition), Strategic_Midcost (including trials where benefactor's cost = 10, 12, or 14 in Strategic condition), Strategic_Highcost (including trials where benefactor's cost = 16, 18, or 20 in Strategic condition) and Altruistic_condition (including all trials the Altruistic condition). Other regressors were the same as GLM 1. A parametric contrast was defined to identify regions that monotonically increased in the strategic condition relative to the altruistic condition (the weights of Strategic_Lowcost, Strategic_Midcost, Strategic_Highcost, and Altruistic_condition were defined as +1, +2, +3, -6 separately) [34].

For all GLMs, events in each regressor were convolved with the canonical hemodynamic response function. Second-level models were constructed as one-sample *t* tests using contrast images from the first-level models. For whole brain analyses, all results were corrected for multiple comparisons with the threshold of voxel-level $p < 0.001$ (uncorrected) combined with cluster-level threshold $p < 0.05$ (FWE-corrected). This threshold provided an acceptable family error control [39,40].

SI Results

Using PCs for ratings on appraisals and emotions as predictors for behaviors

By plotting the difference in the monetary allocation between Strategic condition and Altruistic condition for each participant, large individual difference was observed in the effect of benefactor's intention on reciprocity after receiving help (Fig. S5A). A group of participants reciprocated more to the benefactor when the benefactor had low expectation for repay, whereas another group of participants reciprocated more when the benefactor was perceived to expect repayment. One possibility for the observed individual differences in beneficiaries' behaviors is that these two groups of people had different intensities in emotional responses after receiving help. However, we found that, when splitting participants into two groups according to the reciprocity difference between Strategic and Altruistic conditions, no significant differences were observed in the change of two appraisal-emotional components or any emotional response across the two experimental conditions between these two groups of participants, which excluded this possibility (Table S12). An alternative explanation was that these two groups of individuals had different weights on two appraisal-emotional motivations when making decisions regarding receiving help. To test this explanation possibility, we conducted linear mixed model for reciprocity by including the principal components for communal (PC1) and obligation (PC2) motivations as fixed effects with by-participant random slopes for each fixed effect. Results demonstrated that both communal and obligation motivations contributed significantly to reciprocity (PC1: $\beta = 0.53 \pm 0.03$, $t = 15.77$; PC2: $\beta = 0.18 \pm 0.03$, $t = 7.03$), while the difference between weights on two motivations varied largely across participants (Mean $= 0.35 \pm 0.03$, [-0.44, 1.24], Fig. S5B). Compared with participants who reciprocated more in the Altruistic condition (N = 54), participants who reciprocated more in the Strategic condition (N = 54) had higher relative weight on obligation (i.e., $\beta_{Obligation} / (|\beta_{Obligation}| + |\beta_{Communal}|)$) ($t_{106} = 4.80$, $p < 0.001$, Fig. S5C).

Similar to reciprocity after receiving help, large individual differences were observed in the effect of benefactor's intention on rejection rate (Fig. S5D). A group of participants rejected more help when the benefactor had low expectation for repay, whereas another group of participants rejected more help when the benefactor was perceived to expect repayment. Compared with the participants who rejected more help in the Altruistic condition (N = 49), participants who rejected more help in the Strategic condition (N = 59) had higher relative weight on exchange motivations ($t_{70.92}$ = 3.91, p < 0.001, Fig. S5, E-F). Consistent with model-based results, these model-free results indicated that, the observed individual differences in beneficiaries' behavior regarding receiving help might be derived from participants' different weights on concerns about communal motivation and obligation when making decisions.

Results of computational modeling

Reciprocity. We examined whether our model of dual motivations (Model 1.1) was capable to capture the variations in participants' amount of reciprocity by fitting it to each participant's trials in which they had to accept the co-player's help (see *SI Methods*, *Computational modeling*). This model explained the variations in reciprocity quite well (r^2 = 0.81, p < 0.001; Fig. 4C), and predicted participants' reciprocity better than a range of alternative models (Table S5), such as: (a) models with different mathematical representations of $U_{Communal}$ and $U_{Obligation}$ (Model 1.2 and Model 1.3), (b) models that only include a single motivation term (Model 1.4 and Model 1.5), (c) models with separate parameters for each term (Model 1.6 and Model 1.7), (d) a model that assumes participants reciprocate as a function of the cost to the benefactor (Model 1.8), and (e) a model that assumes that participants are motivated to minimize inequity in payments [22] (Model 1.9). Parameter recovery tests indicated that the parameters of Model 1.1 were identifiable (r = 0.94 ± 0.07, p < 0.001; Table S7). The individuals' relative weights on obligation motivation estimated from our computational modeling were positively correlated with those estimated from linear mixed model, which used the two PCs for participants' actual ratings on appraisals and emotions (Fig. 3E) as predictors for the amount of reciprocity (r = 0.61, p <

0.001).

A simulation of the model across varying combinations of the Θ, Φ and κ parameters reveals diverging predictions of the beneficiaries' response to altruistic and strategic favors (Fig. 4E). Not surprisingly, greedier individuals (higher Θ) were less likely to reciprocate others' favors. However, reciprocity changes as a function of the tradeoff between communal (Φ) and obligation (1 - Φ) motivations and interacts with the intention inference parameter (κ). As the emphasis on obligation increased, the amount of reciprocity to strategic favors increased whereas that to altruistic favors decreased; this effect was enhanced as κ increased. We found that most participants had low Θ values (i.e., greed), but showed a wide range of individual differences in κ and Φ parameters (Fig. 4F). Interestingly, the degree to which the perceived strategic intention reduced the perceived altruistic intention during intention inference (κ) was positively associated with the relative weight on obligation (1-Φ) during reciprocity (r = 0.79, p < 0.001). Given our model simulations, this correlation did not seem to be driven by the intrinsic non-dependence between parameters, and participants with high relative weight on obligation and low κ or low relative weight on obligation and high κ were also observed in the current study. This result suggests that the participants who cared more about the benefactor's strategic intention during intention inference also tended to be motivated by obligation when deciding how much money to reciprocate. Moreover, individual variation in relative weight on communal motivation (Φ) was positively associated with self-reported tendencies to apply a communal norm in relative to exchange norm in everyday reciprocity, as measured by the communal orientation scale and exchange orientation scale [6,7] (see Additional measures; Study 2, r = 0.23, p = 0.019; Study 3, r = 0.37, p = 0.007). These results provide another level of validation supporting that our model was accurately capturing our predicted psychological constructs.

Decision of whether to accept help. We examined the variation in rejection rate exhibited by participants in the task by fitting the dual motivational model to each participant's trials in which they choose whether to accept help (see *SI Methods, Computational modeling*). The softmax function for binary choice add one more parameter (i.e., the temperature parameter λ) to estimate, which may reduce the quality of model fitting due to the limited number of trials. Therefore, in addition to

the model with κ as a free parameter (Mode 2.1), we also included a model (Model 2.4) that applied individual's κ obtained from the best model for reciprocity to reduce the number of parameters. Model comparison demonstrated that the model with reciprocity κ (Model 2.4) was favored over the model with κ as a free parameter (Model 2.1) and a range of alternative models (Table S6), such as: (a) models that only include a single motivation term (Model 2.2, Model 2.3 and Model 2.5), (b) models with separate parameters for each term (Model 2.6 and Model 2.7), and (c) a model that assumes that participants are motivated to minimize inequity in payments [22] (Model 2.9). Moreover, the model with reciprocity κ and Φ ranged from -1 to 1 (Model 2.4) performed better than the model with reciprocity κ and Φ ranged from 0 to 1 (Model 2.8). This model with reciprocity κ and with Φ ranged from -1 to 1 (Model 2.4) correctly predicted 80.00% of participants' choices and captured the behavioral pattern in which individuals rejected more help in Strategic conditions than in Altruistic conditions (Fig. 4D). Parameter recovery exercises demonstrated that we were able to recover the parameters used to simulate data in Model 2.4 with a relatively high accuracy ($r = 0.67 \pm 0.36$, $p < 0.001$) compared with other models (Table S8). These results supported our hypothesis that individuals make trade-off between communal and obligation motivations to decide whether to accept help and how much to reciprocate after receiving help. The individual's relative weight on obligation estimated from our computational modeling was positively correlated with that estimated from linear mixed model, which used the two PCs for participants' actual ratings on appraisals and emotions (Fig. 3E) as predictors for decisions of whether reject help ($r = 0.31$, $p < 0.001$).

To be noted, the parameter κ represented to what extent the second-order belief of benefactor's expectation for repayment reduced the participant's perceived care during intention inference. Therefore, the result that the model with reciprocity κ (Model 2.4) was favored over the model with κ as a free parameter (Model 2.1) indicated a potentially shared mechanism underlying the intention inference processes behind the decisions of whether to accept help and the reciprocal behaviors. In contrast, no significant correlation was observed between parameters related to the trade-off between different motivations (Θ and Φ) estimated from decisions of

31

whether to accept help and those estimated from the amount of reciprocity (Θ: $r = 0.12$, $p = 0.204$; Φ: $r = -0.03$, $p = 0.738$). This may indicate that there existed potentially differential mechanisms underlying the motivation trade-off processing during reciprocity and during the decisions of whether to accept help.

Reference

1   Ellsworth, P. C. & Scherer, K. R. Appraisal processes in emotion. *Handbook of affective sciences* **572**, V595 (2003).

2   Frijda, N. H. The Place of Appraisal in Emotion. *Cognition & Emotion* **7**, 357-387 (1993).

3   Lazarus, R. S. & Smith, C. A. Knowledge and appraisal in the cognition—emotion relationship. *Cognition & Emotion* **2**, 281-300 (1988).

4   Frijda, N. H., Kuipers, P. & Ter Schure, E. Relations among emotion, appraisal, and emotional action readiness. *J Pers Soc Psychol* **57**, 212 (1989).

5   Inui, K., Tran, T. D., Hoshiyama, M. & Kakigi, R. Preferential stimulation of Adelta fibers by intra-epidermal needle electrode in humans. *Pain* **96**, 247-252 (2002).

6   Clark, M. S., Oullette, R., Powell, M. C. & Milberg, S. Recipient's mood, relationship type, and helping. *J Pers Soc Psychol* **53**, 94 (1987).

7   Mills, J. & Clark, M. S. in *Theoretical frameworks for personal relationships.* 29-42 (Lawrence Erlbaum Associates, Inc, 1994).

8   Racine, J. S. RStudio: A Platform‐Independent IDE for R and Sweave. *J. Appl. Econometrics* **27**, 167-172 (2012).

9   Pérez, F. & Granger, B. E. IPython: a system for interactive scientific computing. *Comput Sci Eng* **9** (2007).

10   Burnham, K. P. & Anderson, D. R. Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research* **33**, 261-304 (2004).

11   Neto, J. L., Santos, A. D., Kaestner, C. A., Alexandre, N. & Santos, D. Document clustering and text summarization.   (2000).

12   Salton, G. & Buckley, C. Term-weighting approaches in automatic text retrieval. *Inform Process Manag* **24**, 513-523 (1988).

13   Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent dirichlet allocation. *J Mach Learn Res* **3**, 993-1022 (2003).

14   Blei, D. M. & Lafferty, J. D. Dynamic topic models. in *Proceedings of the 23rd international conference on Machine learning.*   113-120 (ACM).

15   Liu, Q. A novel Chinese text topic extraction method based on LDA. in *International Conference on Computer Science & Network Technology.* (2016).

16   Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823* (2014).

17   Blair, R. J. The neurobiology of psychopathic traits in youths. *Nat Rev Neurosci* **14**, 786-799 (2013).

18   Raiche, G., Magis, D. & Raiche, M. G. Package 'nFactors'. *Parallel Analysis and Non Graphical Solutions to the Cattell Scree Test. Available at CRAN repository: http://cran. r-project. org/web/packages/n Factors/n Factors. pdf* (2010).

19   Revelle, W. An overview of the psych package. *Department of Psychology*

Northwestern University. Accessed on March **3**, 2012 (2011).

20 Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods* **4**, 272 (1999).

21 Dufwenberg, M. & Kirchsteiger, G. A theory of sequential reciprocity. *Games Econ. Behav.* **47**, 268-298 (2004).

22 Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Quart J Econ* **114**, 817-868 (1999).

23 Rabin, M. Incorporating fairness into game theory and economics. *The American economic review*, 1281-1302 (1993).

24 Baumeister, R. F., Stillwell, A. M. & Heatherton, T. F. Guilt: an interpersonal approach. *Psychol Bull* **115**, 243-267 (1994).

25 Greenberg, M. S. in *Social exchange* 3-26 (Springer, 1980).

26 Greenberg, M. S. & Westcott, D. R. Indebtedness as a mediator of reactions to aid. *New directions in helping* **1**, 85-112 (1983).

27 Elfers, J. & Hlava, P. *The Spectrum of Gratitude Experience.* (Springer, 2016).

28 McCullough, M. E., Kilpatrick, S. D., Emmons, R. A. & Larson, D. B. Is gratitude a moral affect? *Psychol Bull* **127**, 249 (2001).

29 Yu, H., Hu, J., Hu, L. & Zhou, X. The voice of conscience: neural bases of interpersonal guilt and compensation. *Soc Cogn Affect Neurosci* **9**, 1150-1158 (2014).

30 Greenberg, M. S. & Shapiro, S. P. Indebtedness: An adverse aspect of asking for and receiving help. *Sociometry*, 290-301 (1971).

31 Algoe, S. B. Find, remind, and bind: The functions of gratitude in everyday relationships. *Social and Personality Psychology Compass* **6**, 455-469 (2012).

32 Algoe, S. B., Haidt, J. & Gable, S. L. Beyond reciprocity: gratitude and relationships in everyday life. *Emotion* **8**, 425 (2008).

33 Battigalli, P. & Dufwenberg, M. Dynamic psychological games. *J. Econ. Theory* **144**, 1-35 (2009).

34 Chang, L. J., Smith, A., Dufwenberg, M. & Sanfey, A. G. Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron* **70**, 560-572 (2011).

35 Sheather, S. *A modern approach to regression with R.* (Springer Science & Business Media, 2009).

36 Bolton, G. E. & Ockenfels, A. ERC: A theory of equity, reciprocity, and competition. *Amer. Econ. Rev.*, 166-193 (2000).

37 Fareri, D. S., Chang, L. J. & Delgado, M. R. Computational substrates of social value in interpersonal collaboration. *J Neurosci* **35**, 8170-8180 (2015).

38 O'doherty, J. P., Hampton, A. & Kim, H. Model‐based fMRI and its application to reward learning and decision making. *Ann Ny Acad Sci* **1104**, 35-53 (2007).

39 Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad*

Sci U S A **113**, 7900-7905 (2016).

40   Flandin, G. & Friston, K. J. Analysis of family‑wise error rates in statistical parametric mapping using random field theory. *Hum Brain Mapp* **hbm.23839** (2017).

Supporting Tables

Table S1. Model comparison for indebtedness rating

	Model	Predictor	Df	AIC	Term	Beta	SE	*t*	*p*
Study 1	Before controlling for benefactor's cost, participant's benefit and social distance								
	Model 1	Obligation	3	5412.8	Obligation	0.34	0.02	16.05	< 0.001
	Model 2	Guilt	3	4239.1	Guilt	0.71	0.02	45.41	< 0.001
	Model 3	Obligation	4	4235.8	Obligation	0.40	0.02	2.31	0.021
		+ Guilt			Guilt	0.70	0.02	40.08	< 0.001
	After controlling for benefactor's cost, participant's benefit and social distance								
	Model 1	Obligation	3	19004.3	Obligation	0.42	0.03	15.87	< 0.001
	Model 2	Guilt	3	17913.3	Guilt	0.74	0.02	43.44	< 0.001
	Model 3	Obligation	4	17907.6	Obligation	0.61	0.02	2.77	0.005
		+ Guilt			Guilt	0.71	0.02	38.26	< 0.001
Study 2	Before controlling for experimental variables								
	Model 1	Obligation	6	6270.1	Obligation	0.35	0.05	6.51	< 0.001
	Model 2	Guilt	6	5312.0	Guilt	0.75	0.03	22.54	< 0.001
	Model 3	Obligation	10	4720.3	Obligation	0.27	0.03	8.04	< 0.001
		+ Guilt			Guilt	0.68	0.04	19.36	< 0.001
	After controlling for experimental variables								
	Model 1	Obligation	6	7744.6	Obligation	0.22	0.04	6.50	< 0.001
	Model 2	Guilt	6	7590.1	Guilt	0.34	0.03	11.30	< 0.001
	Model 3	Obligation	10	7503.1	Obligation	0.17	0.03	10.23	< 0.001
		+ Guilt			Guilt	0.30	0.03	6.24	< 0.001
Study 3	Before controlling for experimental variables								
	Model 1	Obligation	6	2319.2	Obligation	0.29	0.07	4.31	< 0.001
	Model 2	Guilt	6	1985.2	Guilt	0.63	0.04	15.83	< 0.001
	Model 3	Obligation	10	1864.8	Obligation	0.20	0.05	3.82	< 0.001
		+ Guilt			Guilt	0.62	0.06	13.64	< 0.001
	After controlling for experimental variables								
	Model 1	Obligation	6	2685.2	Obligation	0.17	0.04	4.36	< 0.001
	Model 2	Guilt	6	2658.8	Guilt	0.22	0.04	5.33	< 0.001
	Model 3	Obligation	10	2649.6	Obligation	0.13	0.04	3.70	< 0.001
		+ Guilt			Guilt	0.20	0.04	5.02	< 0.001

Table S2. Experimental designs for Study 2 and Study 3

Study	Sample size	Trial number	Experimental variables	Cost manipulation	Efficiency manipulation
Study 2a	51	48	Extra information about benefactor's intention, and benefactor's cost	5, 7, 8, 9, 10, 11, 12,14, 15, 16, 18, 20	1
Study 2b	57	56	Extra information about benefactor's intention, benefactor's cost, and efficiency	4, 8, 12, 16, 20	0.5, 1, 1.5
Study 3	53	54	Extra information about benefactor's intention, and benefactor's cost	4, 6, 8, 10, 12, 14, 16, 18, 20	1

Table S3-1. The effects of the extra information about benefactor's intention and benefactor's cost on participants' emotional and behavioral responses (Study 2a)

Dependent variable	Predictors	Beta	SE	*t* (*z*)	*p*
Second-order belief	Benefactor's cost	0.37	0.03	13.78	< 0.001
	Extra information about benefactor's intention	0.56	0.05	11.41	< 0.001
	Benefactor's cost×Extra information	0.18	0.02	7.38	< 0.001
Perceived care	Benefactor's cost	0.51	0.04	14.57	< 0.001
	Extra information about benefactor's intention	-0.38	0.04	-9.80	< 0.001
	Benefactor's cost×Extra information	-0.08	0.02	-3.99	< 0.001
Gratitude	Benefactor's cost	0.59	0.04	12.64	< 0.001
	Extra information about benefactor's intention	-0.34	0.04	-9.35	< 0.001
	Benefactor's cost×Extra information	-0.05	0.02	-2.54	0.014
Indebtedness	Benefactor's cost	0.33	0.03	11.26	< 0.001
	Extra information about benefactor's intention	-0.27	0.04	-6.46	< 0.001
	Benefactor's cost×Extra information	-0.30	0.02	-1.56	0.126
Guilt	Benefactor's cost	0.43	0.03	12.31	< 0.001
	Extra information about benefactor's intention	-0.07	0.06	-1.10	0.277
	Benefactor's cost×Extra information	0.00	0.02	0.08	0.935
Obligation	Benefactor's cost	0.13	0.04	3.67	< 0.001
	Extra information about benefactor's intention	0.36	0.05	6.91	< 0.001
	Benefactor's cost×Extra information	0.17	0.02	6.22	< 0.001
Reciprocity	Benefactor's cost	0.64	0.03	22.92	< 0.001
	Extra information about benefactor's intention	-0.05	0.03	-1.66	0.103
	Benefactor's cost×Extra information	-0.03	0.02	-1.47	0.147
Decision to reject help	Benefactor's cost	-0.32	0.21	-1.56	0.120
	Extra information about benefactor's intention	0.30	0.10	3.07	< 0.001

Table S3-2. The effects of the extra information about benefactor's intention, benefactor's cost, and efficiency on participants' emotional and behavioral responses (Study 2b)

Dependent variable	Predictors	Beta	SE	*t* (*z*)	*p*
Second-order belief	Benefactor's cost	0.46	0.03	15.63	< 0.001
	Extra information about benefactor's	0.50	0.05	10.76	< 0.001
	Benefactor's cost×Extra information	0.25	0.02	10.91	< 0.001
Perceived care	Benefactor's cost	0.73	0.04	19.51	< 0.001
	Extra information about benefactor's	-0.25	0.02	-10.55	< 0.001
	Efficiency	0.04	0.01	3.95	< 0.001
	Benefactor's cost×Extra information	-0.08	0.01	-5.34	< 0.001
Gratitude	Benefactor's cost	0.60	0.04	14.72	< 0.001
	Extra information about benefactor's	-0.23	0.02	-9.54	< 0.001
	Efficiency	0.05	0.01	3.37	< 0.001
	Benefactor's cost×Extra information	-0.05	0.02	-3.28	< 0.001
Indebtedness	Benefactor's cost	0.60	0.04	16.25	< 0.001
	Extra information about benefactor's	-0.11	0.02	-4.65	< 0.001
	Efficiency	0.04	0.01	2.54	0.010
	Benefactor's cost×Extra information	-0.02	0.02	-1.35	0.180
Guilt	Benefactor's cost	0.48	0.04	12.80	< 0.001
	Extra information about benefactor's	-0.23	0.03	-8.66	< 0.001
	Benefactor's cost×Extra information	-0.07	0.01	-4.74	< 0.001
Obligation	Benefactor's cost	0.30	0.04	7.27	< 0.001
	Extra information about benefactor's	0.25	0.04	6.25	< 0.001
	Benefactor's cost×Extra information	0.10	0.02	6.46	< 0.001
Reciprocity	Benefactor's cost	0.64	0.04	15.82	< 0.001
	Extra information about benefactor's	-0.05	0.02	-3.25	< 0.001
	Efficiency	0.08	0.02	4.90	< 0.001
	Benefactor's cost×Extra information	-0.04	0.01	-2.98	< 0.001
Decision to reject help	Benefactor's cost	-0.97	0.16	-6.00	< 0.001
	Extra information about benefactor's	0.27	0.10	2.84	< 0.001
	Efficiency	-0.49	0.11	-4.39	< 0.001

Table S3-3. The effects of the extra information about benefactor's intention, benefactor's cost, and efficiency on participants' emotional and behavioral responses (combining data of Studies 2a and 2b)

Dependent variable	Predictors	Beta	SE	*t* (*z*)	*p*
Second-order belief	Benefactor's cost	0.42	0.02	20.84	< 0.001
	Extra information about benefactor's intention	0.53	0.03	15.71	< 0.001
	Benefactor's cost×Extra information	0.22	0.02	13.13	< 0.001
Perceived care	Benefactor's cost	0.63	0.03	23.68	< 0.001
	Extra information about benefactor's intention	-0.31	0.02	-13.90	< 0.001
	Efficiency	0.03	0.01	3.60	< 0.001
	Benefactor's cost×Extra information	-0.08	0.01	-6.65	< 0.001
Gratitude	Benefactor's cost	0.55	0.03	19.36	< 0.001
	Extra information about benefactor's intention	-0.27	0.02	-13.18	< 0.001
	Efficiency	0.04	0.01	3.67	< 0.001
	Benefactor's cost×Extra information	-0.06	0.01	-4.20	< 0.001
Indebtedness	Benefactor's cost	0.52	0.03	20.22	< 0.001
	Extra information about benefactor's intention	-0.09	0.03	-2.98	0.004
	Efficiency	0.02	0.01	2.38	0.020
	Benefactor's cost×Extra information	-0.01	0.01	-0.71	0.481
Guilt	Benefactor's cost	0.41	0.02	17.04	< 0.001
	Extra information about benefactor's intention	-0.25	0.02	-10.30	< 0.001
	Benefactor's cost×Extra information	-0.05	0.01	-4.28	< 0.001
Obligation	Benefactor's cost	0.22	0.03	7.71	< 0.001
	Extra information about benefactor's intention	0.30	0.03	9.28	< 0.001
	Benefactor's cost×Extra information	0.11	0.01	8.85	< 0.001
Reciprocity	Benefactor's cost	0.64	0.02	25.77	< 0.001
	Extra information about benefactor's intention	-0.05	0.02	-3.30	0.001
	Efficiency	0.06	0.01	4.76	< 0.001
	Benefactor's cost×Extra information	-0.03	0.01	-3.02	0.003
Decision to reject help	Benefactor's cost	-0.72	0.16	-1564.3	< 0.001
	Extra information about benefactor's intention	0.28	0.10	617.00	< 0.001
	Efficiency	-0.33	0.11	-715.20	< 0.001

Table S4. The effects of the extra information about benefactor's intention and benefactor's cost on participants' emotional and behavioral responses (fMRI study)

Dependent variable	Predictors	Beta	SE	*t* (*z*)	*p*
Second-order belief	Benefactor's cost	0.44	0.02	22.20	< 0.001
	Extra information about benefactor's intention	0.65	0.04	15.88	< 0.001
	Benefactor's cost×Extra information	0.29	0.02	15.21	< 0.001
Perceived care	Benefactor's cost	0.66	0.03	19.23	< 0.001
	Extra information about benefactor's intention	-0.29	0.03	-10.00	< 0.001
	Benefactor's cost×Extra information	-0.09	0.02	-4.16	< 0.001
Gratitude	Benefactor's cost	0.66	0.03	20.45	< 0.001
	Extra information about benefactor's intention	-0.28	0.03	-9.61	< 0.001
	Benefactor's cost×Extra information	-0.06	0.02	-2.77	0.008
Indebtedness	Benefactor's cost	0.46	0.04	11.89	< 0.001
	Extra information about benefactor's intention	-0.26	0.04	-6.64	< 0.001
	Benefactor's cost×Extra information	-0.09	0.02	-4.05	< 0.001
Guilt	Benefactor's cost	0.56	0.04	15.15	< 0.001
	Extra information about benefactor's intention	-0.11	0.03	-3.45	0.001
	Benefactor's cost×Extra information	0.02	0.02	0.87	0.388
Obligation	Benefactor's cost	0.23	0.05	4.89	< 0.001
	Extra information about benefactor's intention	0.29	0.04	6.67	< 0.001
	Benefactor's cost×Extra information	0.17	0.02	7.53	< 0.001
Reciprocity	Benefactor's cost	0.77	0.03	23.56	< 0.001
	Extra information about benefactor's intention	-0.07	0.02	-4.21	< 0.001
	Benefactor's cost×Extra information	-0.02	0.01	-2.65	0.009

Table S5. Model comparison for reciprocity

Model	Model description	Average Sum of Squared Error			Average AIC		
		Study 2a	Study 2b	Combined	Study 2a	Study 2b	Combined
Model 1.1	**DMM Quadratic version**	**4331.00**	**3938.60**	**4123.90**	**117.51**	**130.49**	**124.36**
Model 1.2	DMM Linear version 1	5550.93	7177.44	17156.56	125.22	148.06	171.16
Model 1.3	DMM Linear version 2	5561.19	7179.69	6409.36	125.29	148.08	137.27
Model 1.4	DMM Quadratic version, no Obli	4773.74	5085.24	6415.40	117.16	136.07	137.32
Model 1.5	DMM Quadratic version, no CM	40365.60	33791.45	4938.14	175.24	190.87	127.14
Model 1.6	DMM Quadratic version, separate parameters, mixture weight of greed	4521.81	4134.60	4317.45	121.34	135.09	128.59
Model 1.7	DMM Quadratic version, separate parameters	4466.52	3840.01	4135.86	121.23	132.08	126.96
Model 1.8	Reciprocity according to benefactor's cost	5475.37	5003.66	36895.91	124.01	133.75	183.49
Model 1.9	Inequity aversion model	13726.42	10080.51	5226.41	150.92	160.58	129.15

Note: DMM, Dual motivational model of indebtedness; CM, communal motivation.

Table S6. Model comparison for decisions of whether to accept help

Model	Model description	Log Likelihood			Average AIC		
		Study 2a	Study 2b	Combined	Study 2a	Study 2b	Combined
Model 2.1	DMM, κ as a free parameter	-206.97	-293.05	-251.24	605.95	810.09	710.93
Model 2.2	DMM, κ free, no Obli	-202.92	-303.73	-254.77	549.85	775.47	665.88
Model 2.3	DMM, κ free, no CM	-237.20	-339.40	-289.76	570.40	790.80	683.75
Model 2.4	**DMM, using reciprocity κ**	**-206.97**	**-297.05**	**-253.30**	**557.95**	**762.10**	**662.94**
Model 2.5	DMM, using reciprocity κ, no Obli	-222.56	-325.51	-275.51	541.12	763.03	655.25
Model 2.6	DMM, using reciprocity κ, separate parameters, mixture weight of greed	-206.97	-297.05	-253.30	605.95	818.10	715.06
Model 2.7	DMM, using reciprocity κ, separate parameters	-206.97	-297.05	-253.30	605.95	818.10	715.06
Model 2.8	DMM, using reciprocity κ, $\Phi \in [0,1]$	-230.03	-320.65	-276.63	604.06	809.30	709.61
Model 2.9	Inequity aversion model	-206.97	-297.66	-253.61	605.95	819.31	715.68

Note: DMM, Dual motivational model of indebtedness; CM, communal motivation.

Table S7. Parameter recovery for reciprocity

Model	Model description	Study 2a		Study 2b		Total	
		$r \pm SE$	p	$r \pm SE$	p	$r \pm SE$	p
Model 1.1	**DMM Quadratic version**	**.93±.07**	<0.001	**.94±.08**	<0.001	**.94±.07**	<0.001
Model 1.2	DMM Linear version 1	.29±.20	<0.001	.36±.20	<0.001	.33±.20	<0.001
Model 1.3	DMM Linear version 2	.23±.16	<0.001	.35±.19	<0.001	.29±.18	<0.001
Model 1.4	DMM Quadratic version, no Obli	.99±.02	<0.001	.90±.09	<0.001	.93±.06	<0.001
Model 1.5	DMM Quadratic version, no CM	.99±.02	<0.001	.98±.04	<0.001	.99±.03	<0.001
Model 1.6	DMM Quadratic version, separate parameters, mixture weight of greed	.56±.28	<0.001	.65±.25	<0.001	.61±.27	<0.001
Model 1.7	DMM Quadratic version, separate parameters	.53±.27	<0.001	.70±.23	<0.001	.62±.25	<0.001
Model 1.8	Reciprocity according to benefactor's cost	.80±.21	<0.001	.82±.19	<0.001	.81±.20	<0.001
Model 1.9	Inequity aversion model	.76±.22	<0.001	.73±.24	<0.001	.74±.23	<0.001

Note: DMM, Dual motivational model of indebtedness; CM, communal motivation.

Table S8. Parameter recovery for decisions of whether to accept help

Model	Model description	Study 2a		Study 2b		Total	
		$r \pm SE$	p	$r \pm SE$	p	$r \pm SE$	p
Model 2.1	DMM, κ as a free parameter	.45±.37	<0.001	.41±.38	<0.001	.43±.40	<0.001
Model 2.2	DMM, κ free, no Obli	.73±.43	<0.001	.50±.48	<0.001	.62±.46	<0.001
Model 2.3	DMM, κ free, no CM	-.05±6.94	0.604	.29±.67	0.002	-.02±4.80	0.738
Model 2.4	**DMM, using reciprocity κ**	**.70±.33**	<0.001	**.65±.39**	<0.001	**.67±.36**	<0.001
Model 2.5	DMM, using reciprocity κ, no Obli	.12±6.23	0.250	.18±1.19	0.059	.08±4.39	0.245
Model 2.6	DMM Quadratic version, separate parameters, mixture weight of greed	.64±.32	<0.001	.62±.28	<0.001	.63±.30	<0.001
Model 2.7	DMM Quadratic version, separate parameters	.42±.36	<0.001	.44±.35	<0.001	.43±.35	<0.001
Model 2.8	DMM, using reciprocity κ, $\Phi \in [0,1]$	.73±3.90	<0.001	.40±.38	<0.001	.73±.2.69	<0.001
Model 2.9	Inequity aversion model	.04±2.15	0.563	.20±.97	0.002	.08±.1.64	0.084

Note: DMM, Dual motivational model of indebtedness; CM, communal motivation.

Table S9. Model estimated parameters for reciprocity

Parameters	Study 2a		Study 2b		Total	
	Mean	SE	Mean	SE	Mean	SE
Θ	0.06	0.01	0.10	0.02	0.08	0.01
W_{CM} （Φ）	0.83	0.03	0.75	0.02	0.79	0.02
W_{EM} （1 - Φ）	0.17	0.03	0.25	0.02	0.21	0.02
κ	0.21	0.03	0.41	0.04	0.32	0.01

Table S10. Model estimated parameters for decisions of whether to accept help

Parameters	Study 2		Study 3		Total	
	Mean	SE	Mean	Parameters	Mean	SE
Θ	0.40	0.03	0.41	0.04	0.41	0.03
W_{CM} （Φ）	-0.16	0.08	0.05	0.10	-0.06	0.07
W_{EM} （1 - $|\Phi|$）	0.84	0.08	0.95	0.10	0.42	0.04

Table S11. Results of whole-brain analysis of fMRI data

Regions	Hemisphere	t	Cluster size (voxels)	MNI coordinates		
				x	y	z
Regions responded parametrically to the amount of reciprocity						
Left dlPFC	L	5.93	209	-45	5	29
Right dlPFC	R	4.91	138	45	11	35
	R	4.65	67	45	35	11
Left IPL	L	4.25	48	-54	-40	53
Right IPL	R	4.80	130	51	-28	47
Precuneus-MOG-ITG	R	6.59	958	51	-52	-13
MOG	L	5.34	637	-30	-67	29
ITG	L	5.30	479	-45	-61	-13
Cerebellum	L	38	3.99	-27	-61	-34
Cerebellum	R	45	4.55	6	-31	-22
Regions responded parametrically to communal motivation (ω_B)						
vmPFC	-	4.19	41	0	35	-22
aINS	L	4.19	23	-24	11	-19
Left dlPFC	L	4.35	79	-48	20	26
	L	4.60	75	-24	29	56
Right dlPFC	R	5.24	251	45	11	38
Precuneus	R	4.46	683	3	-46	38
ITG	R	4.85	128	48	-46	-16
ITG	L	4.64	198	-54	-76	-7
MOG	L	4.89	490	-30	-76	32
Calcarine	R	4.29	61	9	-91	5
Regions identified in parametric contrast for obligation motivation (E_B'')						
dmPFC	L	4.39	31	-9	44	41
Left TPJ	L	3.86	42	-57	-61	26

Note: vmPFC = ventromedial prefrontal cortex; aINS = anterior insula; IPL = inferior parietal lobule; dlPFC = dorsolateral prefrontal cortex; ITG = Inferior temporal gyrus; dmPFC = dorsomedial prefrontal lobe; dACC = dorsal anterior cingulate gyrus; MOG = middle occipital gyrus. For whole-brain analyses, all results were corrected for multiple comparisons using the threshold of voxel-level $P < 0.001$ (uncorrected) combined with cluster-level threshold $P < 0.05$ [family-wise error (FWE) corrected].

Table S12. The appraisal-emotional responses for participants with different behavioral patterns

| Reciprocity after receiving help						
Appraisal-emotional responses (Altruistic vs. Strategic)	Participants reciprocated more in the Altruistic condition		Participants who reciprocated more in the Strategic condition		Independent sample t-test	
	Mean	SE	Mean	SE	t_{106}	p
PC1	35.78	4.63	44.84	4.80	-1.34	0.182
PC2	-18.71	2.94	-16.98	2.21	-0.48	0.634
Gratitude	13.40	1.47	14.31	1.51	-0.43	0.670
Indebtedness	2.40	3.06	7.36	1.97	-1.40	0.164
Guilt	12.22	1.91	14.95	1.87	-1.01	0.313
Afraid	-15.88	2.79	-16.88	2.25	0.17	0.866
Decision to reject help						
Appraisal-emotional responses (Altruistic vs. Strategic)	Participants rejected more help in the Altruistic condition		Participants who rejected more help in the Strategic condition		Independent sample t-test	
	Mean	SE	Mean	SE	t_{106}	p
PC1	43.58	4.17	37.88	5.30	0.84	0.401
PC2	-19.74	2.58	-15.94	2.49	-1.10	0.273
Gratitude	13.71	1.10	14.08	1.81	-0.18	0.861
Indebtedness	6.60	2.41	3.61	2.59	0.85	0.400
Guilt	14.63	1.62	12.80	2.14	0.68	0.497
Afraid	-18.46	2.49	-13.96	2.46	-1.29	0.201
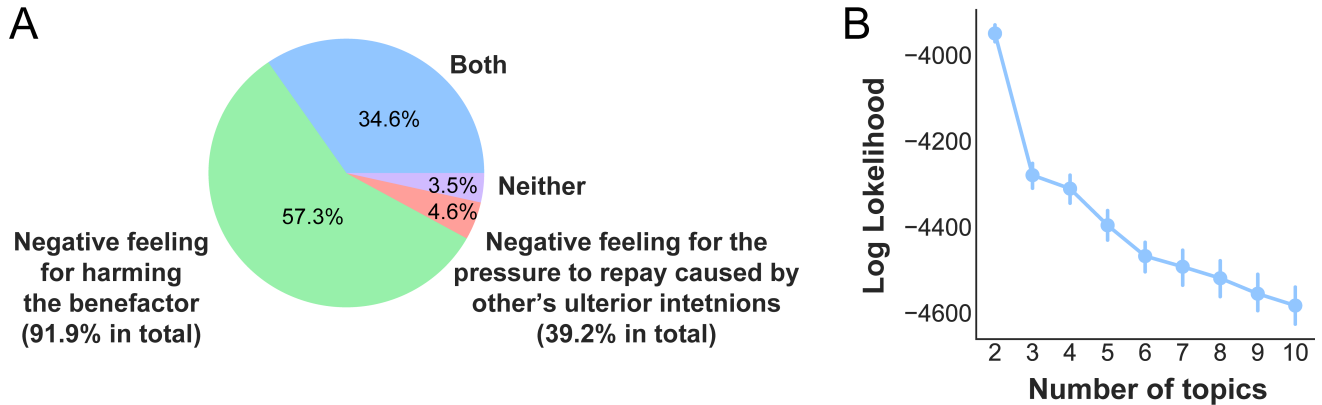
Fig. S1 Definition of indebtedness. (A) The frequency of choosing each option in the question "In daily life, what do you think is/ are the source(s) of indebtedness?" While 57.3% of participants indicated the negative feeling for harming the benefactor as the single source of indebtedness, 4.6% of participants indicated the negative feeling for the pressure of repayment caused by other's ulterior intentions as the single source of indebtedness. 34.6% of participants indicated both types of negative emotions contributing to indebtedness, and 3.5% of participants indicated neither of them as the source of indebtedness. **(B)** Log-likelihood values for Latent Dirichlet Allocation (LDA) based topic modeling with different number of topics. Results demonstrated that the LDA model with 2 topics outperformed models with other number of topics (highest log-likelihood).
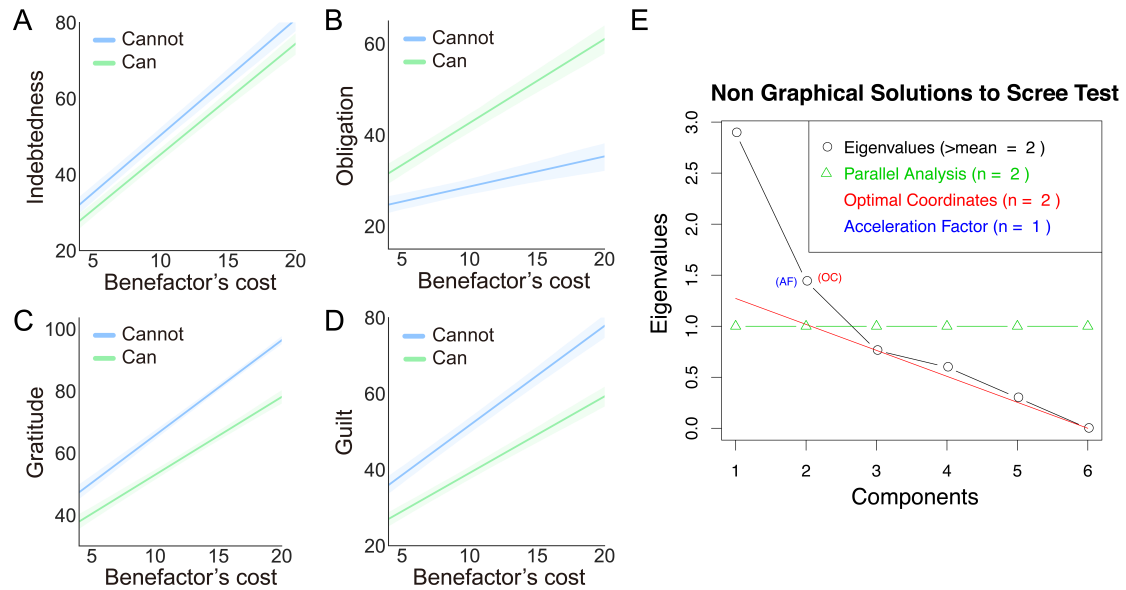
Fig. S2 (A-D) Participant's ratings on indebtedness, the sense of obligation, gratitude and guilt plotted as functions of the extra information about benefactor's intention and benefactor's cost. **(E)** To determine the number of components to retain in the principal component analysis (PCA), the correlation matrix between appraisals and emotions was submitted to a parallel analysis [14]. Parallel analysis performed a principal factor decomposition of the data matrix and compared it to a principal factor decomposition of a randomized data matrix. This analysis yielded components whose eigenvalues (magnitudes) were greater in the observed data relative to the randomized data. The nScree function was used to return an analysis of the number of components to retain using four retention criteria: the Kaiser-Guttman, the scree test (optimal coordinates and acceleration factor), and the parallel analysis. Three of the four criteria pointed to a two-component solution except for the acceleration factor.
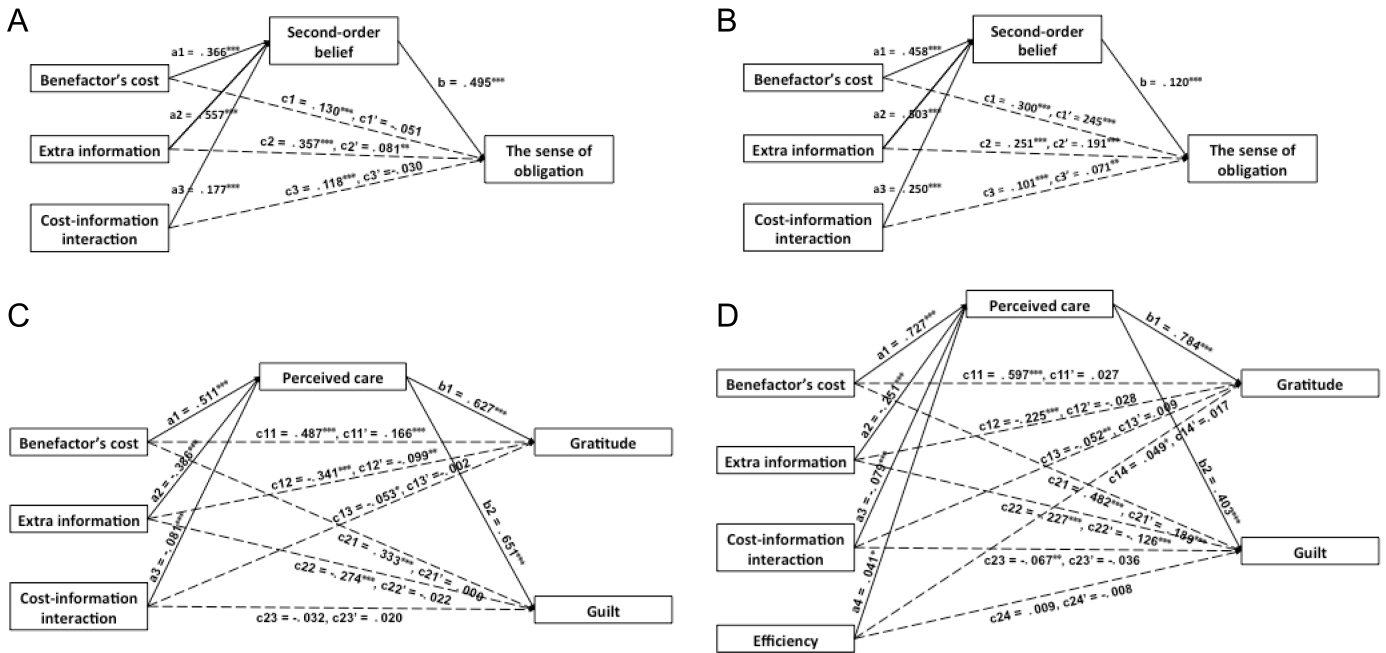
Fig. S3 Results of mediation analyses. (A-B) Second-order belief mediated the effects of experimental variables on obligation. The normalized coefficients of the mediating effects of second-order belief ratings on benefactor's cost to obligation ratings, extra information to obligation ratings, cost-information interaction to obligation ratings were 0.181, $p < 0.001$; 0.275, $p < 0.001$; 0.088, $p < 0.001$, respectively in Study 2a (A), and 0.055, $p < 0.001$; 0.060, $p < 0.001$; 0.030, $p < 0.001$, respectively in Study 2b (B). **(C-D)** Perceived care mediated the effects of experimental variables on gratitude and guilt. The normalized coefficients of the mediating effects of perceived care ratings on benefactor's cost to gratitude ratings, extra information to gratitude ratings, cost-information interaction to gratitude ratings were 0.320, $p < 0.001$; -0.242, $p < 0.001$; -0.051, $p < 0.001$, respectively in Study 2a (C). The normalized coefficients of the mediating effects of perceived care ratings on benefactor's cost to guilt ratings, extra information to guilt ratings, cost-information interaction to guilt ratings were 0.332, $p < 0.001$; -0.251, $p < 0.001$; -0.053, $p < 0.001$, respectively in Study 2b (C). The normalized coefficients of the mediating effects of perceived care ratings on benefactor's cost to gratitude ratings, extra information to gratitude ratings, cost-information interaction to gratitude ratings, efficiency to gratitude ratings were 0.570, $p < 0.001$; -0.197, $p < 0.001$; -0.062, $p < 0.001$; 0.032, $p = 0.012$, respectively in Study 2a (D). The normalized coefficients of the mediating effects of perceived care ratings on benefactor's cost to ratings of guilt, extra information to guilt ratings, cost-information interaction to guilt ratings, efficiency to guilt ratings were 0.293, $p < 0.001$; -0.101, $p < 0.001$; -0.032, $p < 0.001$; 0.016, $p = 0.014$, respectively in Study 2b (D).
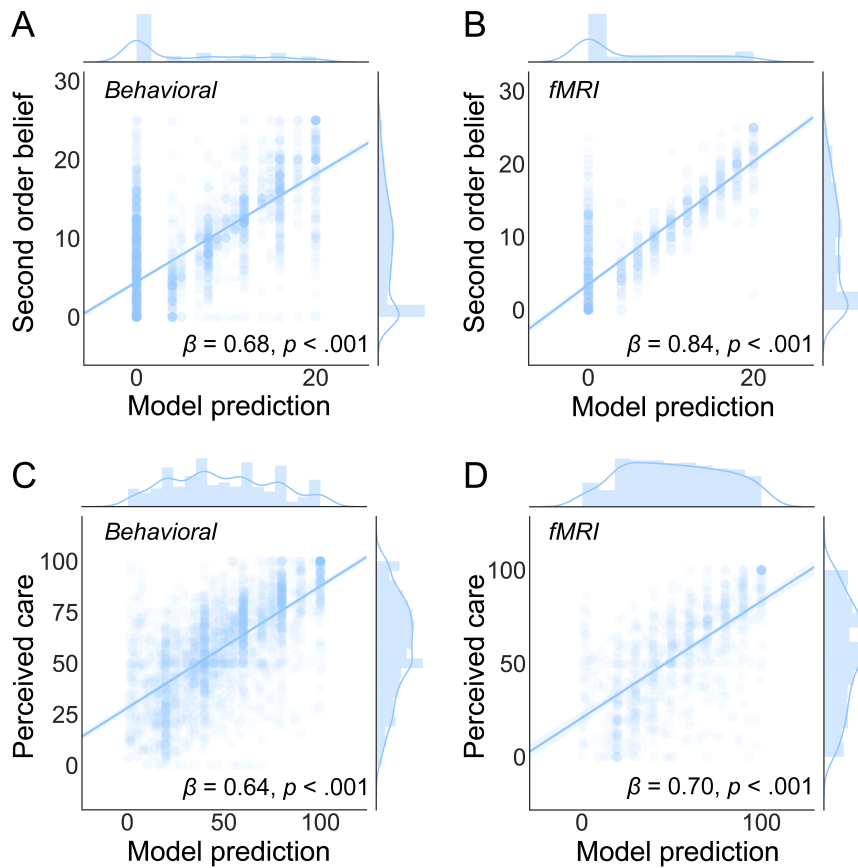
Fig. S4 Model predictions for appraisals on benefactor's intentions. The terms of second-order belief of the benefactor's expectations for repayment (E_B'') and perceived care (ω_B) in our model were able to accurately capture self-reported appraisals of second-order belief ($\beta_{behavioral} = 0.68 \pm 0.03$ (mean ± SE), $t = 21.48$, $p < 0.001$; $\beta_{fMRI} = 0.84 \pm 0.04$, $t = 21.89$, $p < 0.001$, A and B) and perceived care ($\beta_{behavioral} = 0.64 \pm 0.02$, $t = 26.76$, $p < 0.001$; $\beta_{fMRI} = 0.70 \pm 0.04$, $t = 17.78$, $p < 0.001$, C and D), which provided further validation that the model representations were reflecting the intended psychological processes.
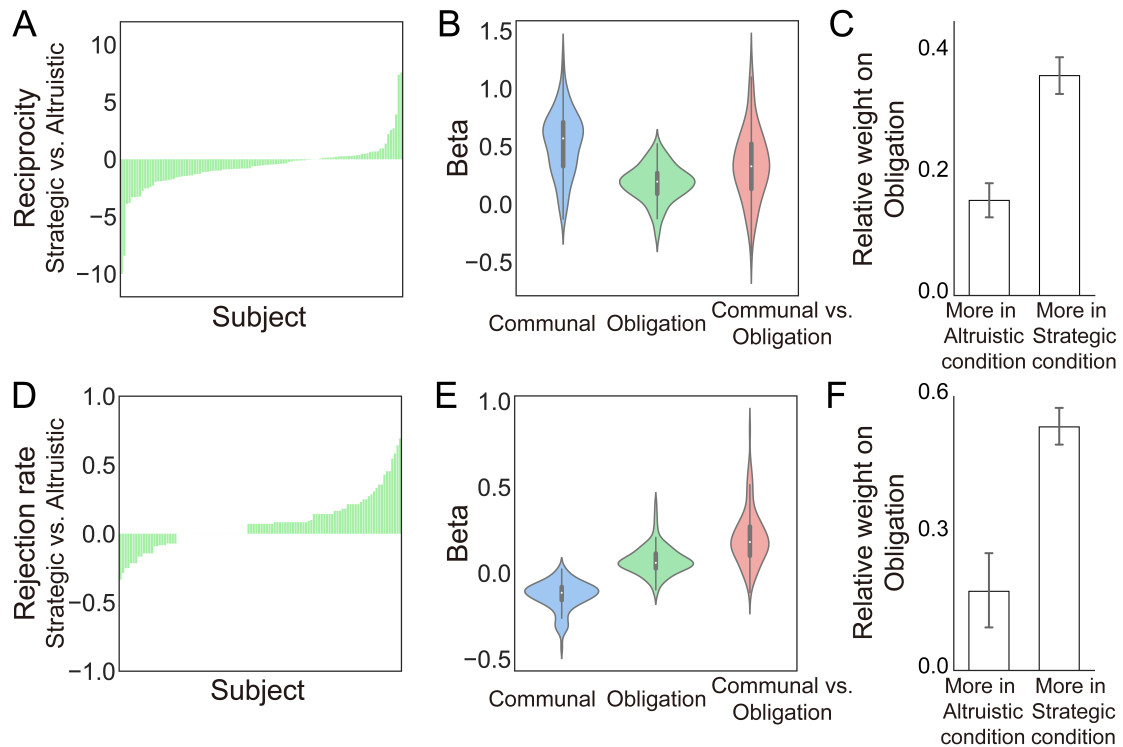
Fig. S5 Using PCs for ratings on appraisals and emotions as predictors for behaviors. (A) The difference in the monetary allocation in Strategic versus Altruistic condition for each participant. **(B)** The contributions of communal motivation (PC1) and obligation motivation (PC2) to reciprocity after receiving help. **(C)** Relative weight on obligation during reciprocity for participants that reciprocated more in Altruistic and Strategic conditions respectively. **(D)** The difference in the frequency to reject help in Strategic versus Altruistic condition for each participant. **(E)** The contributions of communal motivation (PC1) and obligation motivation (PC2) to decisions of whether to reject help. **(F)** Relative weight on obligation when deciding whether to reject help for participants that reciprocated more in Altruistic and Strategic conditions respectively.
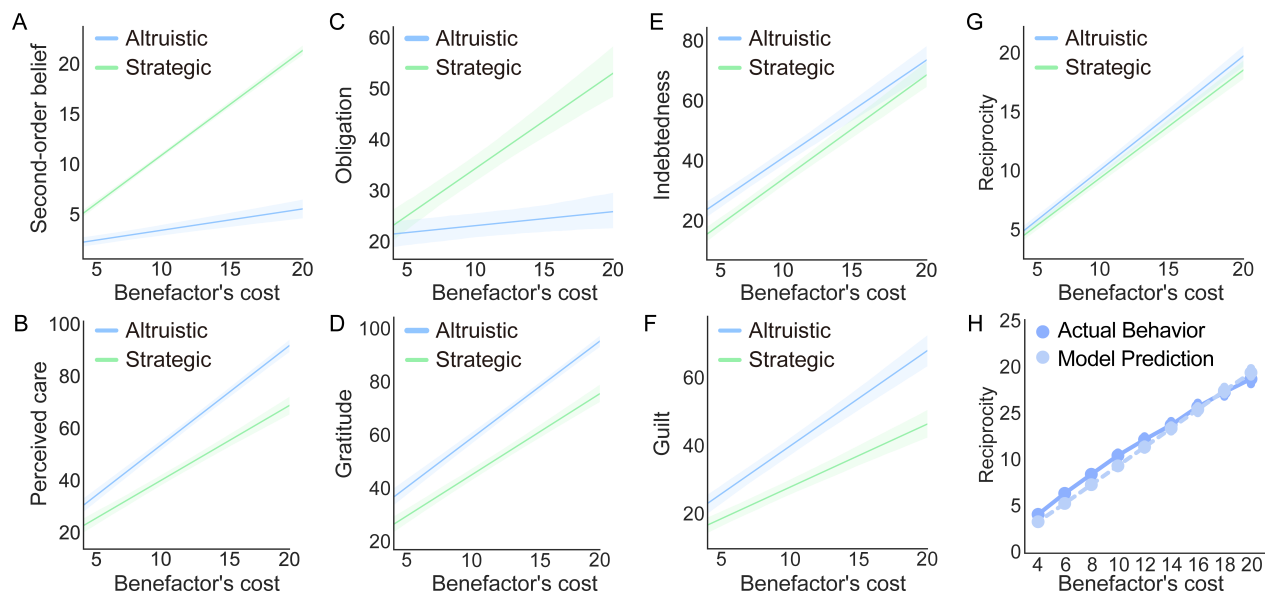
Fig. S6 Appraisal, emotional and behavioral responses in the fMRI study (Study 3) replicated that in the behavioral studies. Compared with the Altruistic condition, the participant's second-order belief of the benefactor's expectation for repay (A), the sense of obligation (C) increased, while the participant's perceived care (B), gratitude (D) and guilt (F) decreased. As the increase of benefactor's cost, the sizes of these effects increased (i.e., significant interaction effects between the extra information about benefactor's intention and benefactor's cost). Participants reported feeling of indebtedness in both conditions, but slightly more in the Altruistic compared to the Strategic condition (E). Participants reciprocated more as the benefactor's increased. This effect was slightly enhanced in the Altruistic relative to the Strategic condition (G). Our computational models accurately captured the patterns of participants' reciprocity after receiving help, $r^2 = 0.95$, $p < 0.001$ (H).

Appendices S1. Online questionnaire for Study 1

Welcome to participate in this questionnaire survey! In the questionnaire, please recall your real life events, and answer the corresponding questions. After answering all the questions, please write a short story about each event. Your story may be used as material for future study. If your story is selected, we will contact you through your contact information and pay you 25 yuan for story authorization. When using the story, we will keep your personal information strictly confidential. Please fill in the answer sheet sincerely and carefully.

To ensure the quality of the data, please check the following box, and promise that you will have at least 15 minutes to conduct the survey and answer each question sincerely. Thank you for your cooperation!

○ I guarantee that I have at least 15 minutes to fill out the questionnaire and answer each question sincerely.

Part 1

Have you received any help in the past one year? [single choice]

○ Yes (Continue)   ○ No (skip Part 1)

Please think carefully about an event in which you received help from others that impressed you the most in the past one year and happened recently.

1. What is the time of this event? (Please select the option that matches the occurrence of the event and is closest to today)

○Within one week ○Within one month,

○Within three months ○Within half a year ○Within one year

2. In this event, did you actively seek help or passively accept help from others? [single choice]

○Actively seek   ○Passive acceptance

3. How was your relationship with this benefactor before receiving help? (0 is very unfamiliar, 100 is very familiar)

4. To what extent were you willing to accept this help? (0 is not willing to accept, 100 very willing to accept)

5. Who was the person that helped you? [single choice]

○Parent ○Sibling ○Spouse/boyfriend/girlfriend ○Other relative ○Friend

○Classmate/Colleague ○Teacher ○Neighbor ○Stranger

○Other (such as service personnel, public servants, etc.. Please fill in the benefactor's specific occupation) _____

6. How helpful was the help? (0 is useless, 100 is very helpful)

7. How much was the benefactor's cost in this help? (0 is not at all, 100 is very big)

8. Before receiving the help, how likely did you think the benefactor would help? (0 is completely impossible, 100 is pretty sure)

9. How grateful did you feel about the benefactor's help?　(0 is not at all, 100 is very strong)

10. How indebted did you feel about the benefactor's help? (0 is not at all, 100 is very strong)

11. How guilty did you feel about the benefactor's help? (0 is not at all, 100 is very strong)

12. How much were you afraid of the benefactor's expectation for repay? (0 is not at all, 100 is very strong)

13. How much pressure did you feel to reciprocate in the future? (0 is not at all, 100 is very strong)

14. To what extent did you think you needed to reciprocate? (0 is not at all, 100 is very strong)

15. To what extent did you think the benefactor expected you to reciprocate? (0 is not at all, 100 is very strong)

16. Compared with the benefit you obtained from the help, how much did you think you needed to reciprocate to the benefactor? (0-50 means less than your benefit, 50 means equal to your benefit, 50-100 means more than your benefit)

17. Compared with the benefactor's cost, how much did you think you needed to reciprocate to the benefactor? (0-50 means less than the benefactor's cost, 50 means equal to the benefactor's cost, 50-100 means more than the benefactor's cost)

18. Compared with the benefit you obtained from the help, how much did you think the benefactor needed you to reciprocate? (0-50 means less than your benefit, 50 means equal to your benefit, 50-100 means more than your benefit)

19. Compared with the benefactor's cost, how much did you think the benefactor needed you to reciprocate? (0-50 means less than the benefactor's cost, 50 means equal to the benefactor's cost, 50-100 means more than the benefactor's cost)

20. Did the benefactor proposed a clear request for reciprocity? [single choice]

○ Yes, please briefly explain the details _____ ○ No

21. To what extent did you think the benefactor cared about your welfare when helping you? (0 is not at all, 100 is very strong)

22. To what extent did you think the benefactor cared about his/her own interests when helping you? (0 is not at all, 100 is very strong)

23. To what extent did you want to reciprocate? (0 is not at all, 100 is very strong)

24. To what extent did you want to repay the favor immediately? (0 is not at all, 100 is very strong)

25. In what way did you want to reciprocate? [multiple choice]

□ Monetary reciprocity □ Help each other

□ Oral thanks   □ Establish cooperative relationship

□ Make friends with him/her □ Gifts □ Other _____

26. Have you reciprocated in some way? [multiple choice]

□ Monetary reciprocity □ Help each other

□ Oral thanks   □ Establish cooperative relationship

□ Make friends with him/her □ Gifts □ Other _____

27. To what extent were you willing to interact or get to know each other further? (0 is not at all, 100 is very strong)

28. Please describe the event in detail in the form of a short story.

(If your story is selected as the material for our experiment, you will receive a story authorization fee of 25 yuan)

Part 2

Have you rejected any help in the past one-year? [single choice]

○ Yes (Continue)   ○ No (skip Part 2)

Please think carefully about an event when you received help from others that impressed you most in the past one year and happened recently.

1. What is the time of this event? (Please select the option that matches the occurrence of the event and is closest to today)

○Within one week ○Within one month,

○Within three months ○Within half a year ○Within one year

2. In this event, did the benefactor actively provide the offer of help or did someone suggest the benefactor to give you help? [single choice]

○Actively provided the offer of help ○Someone suggested the benefactor to help

3. How was your relationship with this benefactor before this event? (0 is very unfamiliar, 100 is very familiar)

4. To what extent did you want to reject the offer of help? (0 is not at all, 100 is very strong)

5. Who was the person that offered to help you? [single choice]

○Parent ○Sibling ○Spouse/boyfriend/girlfriend ○Other relative ○Friend

○Classmate/Colleague ○Teacher ○Neighbor ○Stranger

○Other (such as service personnel, public servants, etc.. Please fill in the benefactor's specific occupation) _____

6. Imagine if you have accepted the help, how helpful would the help be? (0 is useless, 100 is very helpful)

7. Imagine if you have accepted the help, how much would the benefactor's cost be in this help? (0 is not at all, 100 is very big)

8. Before receiving the help, how likely did you think the benefactor would help? (0 is completely impossible, 100 is pretty sure)

9. Imagine if you have accepted the help, how grateful would you feel about the benefactor's help? (0 is not at all, 100 is very strong)

10. Imagine if you have accepted the help, how indebted would you feel about the benefactor's help? (0 is not at all, 100 is very strong)

11. Imagine if you have accepted the help, how guilty would you feel about the benefactor's help? (0 is not at all, 100 is very strong)

12. Imagine if you have accepted the help, how much you were afraid of the benefactor's expectation for repay? (0 is not at all, 100 is very strong)

13. Imagine if you have accepted the help, how much pressure would you feel to reciprocate in the future? (0 is not at all, 100 is very strong)

14. Imagine if you have accepted the help, to what extent did you think you needed to reciprocate? (0 is not at all, 100 is very strong)

15. Imagine if you have accepted the help, to what extent would you think the benefactor expected you to reciprocate? (0 is not at all, 100 is very strong)

16. Imagine if you have accepted the help, compared with the benefit you obtained from the help, how much would you think you needed to reciprocate to the benefactor? (0-50 means less than your benefit, 50 means equal to your benefit, 50-100 means more than your benefit)

17. Imagine if you have accepted the help, compared with the benefactor's cost, how much would you think you needed to reciprocate to the benefactor? (0-50 means less than the benefactor's cost, 50 means equal to the benefactor's cost, 50-100 means more than the benefactor's cost)

18. Imagine if you have accepted the help, compared with the benefit you obtained from the help, how much would you think the benefactor needed you to reciprocate? (0-50 means less than your benefit, 50 means equal to your benefit, 50-100 means more than your benefit)

19. Imagine if you have accepted the help, compared with the benefactor's cost, how much would you think the benefactor needed you to reciprocate? (0-50 means less than the benefactor's cost, 50 means equal to the benefactor's cost, 50-100 means more than the benefactor's cost)

20. Did the benefactor ask for repayment before helping you? [single choice]

○Yes, please briefly explain the details _____   ○No

21. To what extent did you think the benefactor cared about your welfare when he/she offered to help you? (0 is not at all, 100 is very strong)

22. To what extent did you think the benefactor cared about his/her own interests when he/she offered to help you? (0 is not at all, 100 is very strong)

23. Imagine if you have accepted the help, to what extent would you want to reciprocate? (0 is not at all, 100 is very strong)

24. Imagine if you have accepted the help, to what extent did you want to repay the favor immediately? (0 is not at all, 100 is very strong)

25. Imagine if you have accepted the help, in what way did you want to reciprocate? [multiple choice]

□ Monetary reciprocity □ Help each other

□ Oral thanks   □ Establish cooperative relationship

□ Make friends with him/her □ Gifts   □ Other

27. What was/were your reason(s) for refusing the offer? [multiple choice]

□ Thought the benefactor's purpose is not pure

□ The anticipatory repayment was too much

□ Limit your freedom

□ Feeling your self-esteem was hurt

□ The benefit from the help was little

□ The benefactor's cost was too much

28. Please describe the event in detail in the form of a short story.

(If your story is selected as the material for our experiment, you will receive a story authorization fee of 25 yuan)

• In the context of helping and receiving help, what is your definition of gratitude?

• In the context of helping and receiving help, what is your definition of indebtedness?

• In daily life, what do you think is/are the source(s) of indebtedness? (Single choice, the order of the first two options was counterbalanced among participants)

　　Negative feeling for harming the benefactor/for cost that the benefactor has paid for helping you

　　Negative feeling for the pressure to repay caused by other's ulterior intentions (e.g., Expectation for repay)

　　Both of the above

　　Neither of the above

Appendices S2. Classification for words in the definition of indebtedness

Classification	Word	English word	Weight	Frequency to be Classified in each level (%)				
				Appraisal	Emotion	Behavior	Person	Other
Appraisal	损失	Loss	0.075	45.0	7.5	36.3	0.0	11.3
	代价	Cost	0.033	41.3	5.0	17.5	2.5	33.8
	不好	Bad	0.014	55.0	30.0	0.0	1.3	13.8
	受损	Harm	0.012	45.0	6.3	45.0	0.0	3.8
	很大	Great	0.012	51.3	5.0	2.5	2.5	38.8
	不必要	Unnecessary	0.010	46.3	12.5	5.0	1.3	35.0
Emotion	愧疚	Guilt	0.269	1.3	97.5	0.0	0.0	1.3
	内疚	Guilt	0.192	0.0	98.8	0.0	1.3	0.0
	亏欠	Feel indebted	0.154	25.0	46.3	26.3	0.0	2.5
	感觉	Feel	0.120	10.0	66.3	15.0	0.0	8.8
	感到	Feel	0.102	5.0	66.3	20.0	0.0	8.8
	觉得	Feel	0.074	15.0	53.8	17.5	0.0	13.8
	对不起	Feel sorry	0.068	8.8	62.5	15.0	0.0	13.8
	想要	Want to	0.052	6.3	56.3	32.5	1.3	3.8
	不安	Uneasy	0.047	0.0	97.5	1.3	0.0	1.3
	麻烦	Trouble	0.041	26.3	36.3	21.3	2.5	13.8
	难受	Uncomfortable	0.040	0.0	98.8	0.0	1.3	0.0
	负罪感	Guilt	0.034	3.8	93.8	0.0	1.3	1.3
	自责	Guilt	0.033	5.0	85.0	8.8	1.3	0.0
	过意不去	Feel sorry	0.026	2.5	95.0	0.0	1.3	1.3
	有愧	Guilt	0.023	1.3	95.0	2.5	0.0	1.3
	感激	Gratitude	0.022	1.3	86.3	12.5	0.0	0.0
	不好意思	Feel sorry	0.020	1.3	91.3	2.5	3.8	1.3
	抱歉	Feel sorry	0.017	2.5	87.5	8.8	1.3	0.0
	不舒服	Uncomfortable	0.016	5.0	92.5	0.0	1.3	1.3
	心里	In the heart	0.013	3.8	41.3	25.0	3.8	26.3
	压力	Pressure	0.013	10.0	72.5	3.8	1.3	12.5
	情感	Emotion	0.013	7.5	70.0	0.0	0.0	22.5
	内疚感	Guilt	0.013	1.3	96.3	0.0	2.5	0.0
	负担	Burden	0.012	16.3	37.5	33.8	0.0	12.5
	痛苦	Painful	0.011	1.3	95.0	3.8	0.0	0.0
	强烈	Strong	0.010	20.0	57.5	2.5	1.3	18.8
	希望	Want to	0.010	16.3	56.3	17.5	0.0	10.0
	歉疚	Guilt	0.010	1.3	96.3	1.3	1.3	0.0
Behavior	帮助	Help	0.312	3.8	0.0	93.8	2.5	0.0
	伤害	Harm	0.101	15.0	11.3	73.8	0.0	0.0
	付出	Cost	0.091	10.0	3.8	82.5	1.3	2.5
	负债	Be in debt	0.090	22.5	13.8	46.3	2.5	15.0
	回报	Repay	0.068	20.0	3.8	66.3	1.3	8.8

Classification	Word	English word	Weight	Frequency to be Classified in each level (%)				
				Appraisal	Emotion	Behavior	Person	Other
	造成	Cause	0.067	15.0	0.0	78.8	0.0	6.3
	损害	Harm	0.065	22.5	5.0	68.8	0.0	3.8
	受到	Receive	0.046	7.5	17.5	45.0	0.0	30.0
	接受	Receive	0.044	3.8	10.0	82.5	1.3	2.5
	产生	Generate	0.037	7.5	3.8	57.5	1.3	30.0
	补偿	Compensate	0.037	12.5	3.8	83.8	0.0	0.0
	牺牲	Sacrifice	0.037	15.0	2.5	75.0	1.3	6.3
	偿还	Repay	0.034	16.3	2.5	77.5	1.3	2.5
	回馈	Repay	0.022	13.8	1.3	78.8	2.5	3.8
	带来	Bring	0.019	6.3	3.8	77.5	1.3	11.3
	收到	Receive	0.017	2.5	2.5	88.8	1.3	5.0
	需要	Need	0.017	21.3	17.5	32.5	1.3	27.5
	影响	Influence	0.016	26.3	11.3	51.3	0.0	11.3
	弥补	Compensate	0.016	5.0	10.0	83.8	1.3	0.0
	行为	Behavior	0.015	7.5	1.3	68.8	1.3	21.3
	给予	Give	0.013	3.8	0.0	93.8	0.0	2.5
	报答	Repay	0.012	8.8	7.5	81.3	1.3	1.3
	得到	Receive	0.012	6.3	2.5	82.5	1.3	7.5
	付出代价	Pay the price	0.012	17.5	2.5	70.0	1.3	8.8
	我会	I will	0.012	16.3	7.5	47.5	5.0	23.8
	做错	Wrongdoings	0.012	28.8	6.3	57.5	2.5	5.0
	做错事	Wrongdoings	0.011	17.5	5.0	72.5	0.0	5.0
	失去	Loss	0.011	15.0	10.0	67.5	0.0	7.5
	导致	Lead to	0.011	18.8	2.5	57.5	1.3	20.0
Person	别人	Other	0.329	1.3	0.0	1.3	96.3	1.3
	他人	Other	0.218	0.0	0.0	2.5	97.5	0.0
	自己	Self	0.202	0.0	0.0	0.0	98.8	1.3
	对方	Other	0.142	2.5	0.0	0.0	96.3	1.3
	帮助者	Benefactor	0.036	2.5	1.3	2.5	88.8	5.0
	其他人	Other	0.026	1.3	0.0	2.5	95.0	1.3
	自身	Self	0.016	2.5	3.8	0.0	83.8	10.0
Other	利益	Benefit	0.131	36.3	7.5	6.3	2.5	47.5
	内心	In the heart	0.114	3.8	43.8	0.0	2.5	50.0
	因为	Because	0.077	18.8	3.8	5.0	1.3	71.3
	心存	In the heart	0.057	6.3	36.3	2.5	1.3	53.8
	人情	Favor	0.040	17.5	20.0	7.5	2.5	52.5
	东西	Things	0.029	7.5	1.3	2.5	8.8	80.0
	心理	Psychological	0.029	7.5	38.8	1.3	2.5	50.0
	事情	Things	0.028	5.0	0.0	8.8	5.0	81.3
	或者	Or	0.024	7.5	0.0	1.3	1.3	90.0

Classification	Word	English word	Weight	Frequency to be Classified in each level (%)				
				Appraisal	Emotion	Behavior	Person	Other
	为了	In order to	0.023	18.8	2.5	21.3	1.3	56.3
	一种	A kind of	0.021	15.0	0.0	2.5	2.5	80.0
	没有	No	0.020	25.0	0.0	2.5	1.3	71.3
	某件事	Something	0.019	3.8	0.0	6.3	15.0	75.0
	有所	Somewhat	0.017	16.3	8.8	7.5	0.0	67.5
	对于	For	0.016	21.3	2.5	10.0	2.5	63.8
	一些	Some	0.015	11.3	0.0	3.8	6.3	78.8
	什么	What	0.015	7.5	2.5	1.3	2.5	86.3
	应该	Should	0.015	30.0	16.3	10.0	1.3	42.5
	程度	Extent	0.014	33.8	2.5	6.3	5.0	52.5
	由于	Because	0.013	23.8	1.3	5.0	1.3	68.8
	原因	Reason	0.013	25.0	3.8	3.8	0.0	67.5
	某事	Something	0.012	5.0	0.0	5.0	11.3	78.8
	某些	Some	0.012	7.5	0.0	5.0	7.5	80.0
	一定	Certainly	0.012	27.5	7.5	7.5	2.5	55.0
	是否	Whether	0.012	37.5	1.3	2.5	1.3	57.5
	感是	Feel	0.012	10.0	28.8	3.8	0.0	57.5
	心中	In the heart	0.011	3.8	40.0	2.5	3.8	50.0
	无法	Cannot	0.011	25.0	11.3	6.3	0.0	57.5
	道德	Moral	0.011	25.0	16.3	11.3	2.5	45.0
	责任	Responsibility	0.010	25.0	23.8	15.0	3.8	32.5