



Original Article

Cooperation: The roles of interpersonal value and gratitude

Adam Smith^{a,b}, Eric J. Pedersen^{a,c}, Daniel E. Forster^a, Michael E. McCullough^a, Debra Lieberman^{a,*}^a Department of Psychology, University of Miami, P.O. Box 248185, Coral Gables, FL, USA^b Department of Psychology, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan^c Department of Psychology and Neuroscience, University of Colorado, Boulder, 345 UCB Boulder, CO, USA

ARTICLE INFO

Article history:

Initial receipt 8 March 2017

Final revision received 23 August 2017

Keywords:

Cooperation

Gratitude

Welfare trade-off ratio (WTR)

ABSTRACT

Here we examine the roles of interpersonal valuation and gratitude in the formation of cooperative relationships. Building on prior work, we draw on the concept of a welfare tradeoff ratio (WTR), an internally computed index of the extent to which one person values another person's welfare relative to his or her own. We test several predictions regarding the effects of benefit delivery on changes in WTR, gratitude, and subsequent cooperation. We show that benefit delivery by a stranger: (i) raises the beneficiary's valuation of the stranger's welfare, (ii) predicts downstream cooperative behavior by the beneficiary toward the stranger, and (iii) is coincident with beneficiaries' expressions of gratitude. We find evidence that cooperation and gratitude, while both sparked via benefit delivery and both underpinned by estimates of welfare valuation, are nevertheless produced in parallel via different paths. Specifically, the updated magnitude—not the initial magnitude or degree of change—of a beneficiary's WTR toward a stranger predicts the beneficiary's downstream cooperative behavior. By contrast, the extent to which the beneficiary's WTR positively *changes*—and not the initial or updated WTR magnitude—predicts gratitude production, a feature proposed to reinforce the benefactor's actions and foreshadow future cooperative intent on the part of the beneficiary. Taken together, our findings point to the possibility that cooperative behavior might operate via internal estimates of welfare valuation, and that gratitude signals benefit reception and the intent to engage in a cooperative relationship.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Among primates, humans are unusual for their high degree of cooperation with non-kin. We engage in social exchange for mutual benefit; we form fitness-enhancing social groups; we develop relatively long pair bonds with mates; and we forge deep engagements with non-kin in friendships. A dominant explanation for the variety of cooperative relationships observed among non-kin in humans—and non-humans as well—has been reciprocal altruism, also known simply as reciprocity (Boyd & Richerson, 1992; Fehr & Fischbacher, 2003; Gintis, Bowles, Boyd, & Fehr, 2003; Nowak & Sigmund, 2005; Tooby & Cosmides, 1996; Trivers, 1971).

Over the past few decades, much research has focused on identifying the psychological systems governing the development and maintenance of reciprocal relationships (e.g., Cosmides & Tooby, 2005; Hammerstein, 2003; Rand, Ohtsuki, & Nowak, 2009; Schino & Aureli, 2009, 2010; Sussman & Cloninger, 2011). Here, we propose and test a model of how reciprocal relationships in humans might be initiated

and maintained based on the assessment of *interpersonal value* and the emotion of gratitude.

As others have noted, gratitude appears well designed to facilitate the development of reciprocal relationships (Algoe & Haidt, 2009; Bartlett, Condon, Cruz, Baumann, & DeSteno, 2012; DeSteno, Bartlett, Baumann, Williams, & Dickens, 2010; Emmons & McCullough, 2003; Forster, Pedersen, Smith, McCullough, & Lieberman, 2017; Lim, 2012; Nowak & Roch, 2007). Gratitude is typically elicited when an individual delivers a benefit intentionally at some personal cost (Gergen, Ellsworth, Maslach, & Seipel, 1975; McCullough, Kilpatrick, Emmons, & Larson, 2001; McCullough, Tsang, & Emmons, 2004; Tesser, Gatewood, & Driver, 1968; Tsang, 2006). Furthermore, feeling grateful leads to desires to return benefits specifically to a benefactor, yet can extend to novel individuals as well (Bartlett & DeSteno, 2006; Lim, 2012). That is, gratitude has been linked to direct reciprocity—helping a past benefactor—as well as indirect reciprocity—helping a novel individual (e.g., Nowak & Roch, 2007).

Despite a growing body of work, there is still much unknown regarding the information processing systems that regulate both the deployment of gratitude and the downstream facilitation of cooperation. Information-processing models have been helpful in guiding research in various domains such as kin detection, sexual attraction, kin-directed altruism, and pathogen avoidance (e.g., Lieberman, Tooby, & Cosmides,

* Corresponding author at: Department of Psychology, University of Miami, P.O. Box 248185, Coral Gables, FL 33124-0751, USA.

E-mail address: debra@miami.edu (D. Lieberman).

2007; Tybur & Lieberman, 2016). By and large, sketches of information processing pathways enable researchers to more precisely specify and generate predictions regarding the nature of the inputs, algorithms, and outputs associated with a particular functionality.

Here we propose an information processing structure that attempts to explain how benefit delivery influences both gratitude and cooperative behaviors. Building on prior work, we suggest gratitude evolved to motivate the formation and strengthening of mutually beneficial relationships by signaling to another individual that an act that he or she performed had beneficial consequences for the beneficiary. Perceptions of benefit delivery update—in the mind of the beneficiary—an internal estimate of how the benefactor values the beneficiary. Updated estimates, when larger than initial estimates, activate gratitude displays and can lead to an increased motivation on the part of the beneficiary to return benefits to the benefactor. Although our focus here is on the systems producing gratitude and motivating the return of benefits to a benefactor, we also propose that an expression of gratitude updates the benefactor's view of the person expressing gratitude as “someone who cares about my welfare” and can motivate cognitions and behaviors that increase the likelihood the benefactor and beneficiary will develop and maintain a mutually welfare-regarding association – or friendship. Thus there are two interacting systems: one activated in the mind of the beneficiary, which includes perceptions of benefit receipt, gratitude expression, and downstream benefit return; and one activated in the mind of the benefactor, which includes gratitude detection, and decision-making systems of whether to develop a mutually beneficial relationship. Both, we contend, require mechanisms for assessing the value another individual places on the self and for indexing how the self should value another person. Here we explore the possibility that the perception of changes in welfare valuation serves as the proximate mechanism by which benefit delivery influences cooperation and by which gratitude kindles cooperative relationships. Recent work posits an internal representation termed the welfare trade-off ratio (WTR) as one possible way welfare valuation is assessed.

1.1. Internal representations of interpersonal value: the welfare trade-off ratio

A basic requirement for any social species is the ability to estimate and cognitively represent the fitness value of interacting with another individual, as indices of value can then govern decisions regarding cooperative versus competitive effort (Sell, Tooby, & Cosmides, 2009). One recent proposal is that aspects of human social decision-making and behavior are adaptively regulated by an internal regulatory variable termed the *welfare trade-off ratio* (WTR; Delton, 2010; Forster et al., 2017; Lim, 2012; Sznycer, 2010; Tooby & Cosmides, 2008; Tooby, Cosmides, Sell, Lieberman, & Sznycer, 2008). A WTR is posited to be an internal representation of interpersonal value and shares similarities with other conceptions of fitness value (e.g., Roberts, 2005). In accord with previous descriptions (see Sell et al., 2009), we conceive of WTR as an estimate that captures how one individual values the welfare of another individual, and that is continually updated based on interactions. Thus, a WTR can represent how Ego values Individual A, how Individual A values Ego, Ego's perception of how Individual A values Individual B, et cetera.

WTRs are summary variables thought to be computed from multiple fitness-relevant inputs. One such input is genetic relatedness (Griffin & West, 2003; Hamilton, 1964; Lieberman et al., 2007). In humans and non-humans alike, cues that would have been reliable correlates of genetic relatedness in ancestral environments are posited to be used by a kinship estimating system to generate an internal *kinship estimate*, which is then taken as input by a system that computes WTRs (Cosmides & Tooby, 2013). Greater certainty of kinship, as assessed via kinship estimates, translates into larger magnitudes of WTR values, which, in turn, regulate motivations to re-direct resources away from

oneself toward the focal individual, or to incur costs to protect the focal individual from harm.

But genetic relatedness is not the only factor that should influence decisions to cooperate. Even within a class of highly related individuals—full siblings, for instance—benefit delivery can yield different outcomes due to disparities in health, reproductive status, and age. A system—such as a WTR system—capable of integrating these additional fitness relevant dimensions with estimates of genetic relatedness would fare even better on the evolutionary fitness landscape. Thus, WTRs likely take as input multiple fitness relevant factors, for instance, another's positive or negative externalities (McCullough, Pedersen, Tabak, & Carter, 2014; Sell, 2011; Sell et al., 2009; Tooby et al., 2008; Tooby & Cosmides, 1996), past experiences of cooperative or exploitive interactions (Krasnow, Cosmides, Pedersen, & Tooby, 2012; Trivers, 1971), and future interaction opportunities (Delton, Krasnow, Cosmides, & Tooby, 2011; Krasnow, Delton, Tooby, & Cosmides, 2013).

Yet another key factor that should regulate one's WTR toward another person and thus, decisions to cooperate, is the perceived value another holds for the self. To illustrate why another's valuation of the self might have had positive fitness consequences on decisions of whether or not to cooperate, consider two individuals that *only* differ in their propensity to deliver benefits (or impose costs). Differences in how they treat you reveals information regarding how each individual values you and, thus, the likely benefits (or costs) of future interactions with each. A system that integrated the likely payoffs of interacting with an individual based on prior interaction history would have allowed individuals to direct their cooperative efforts to valuable partners and prevented potential exploitation.

In sum, the proposal is that fitness-relevant inputs are integrated by a system that computes a summary variable of how the self should regard another individual, that is, $WTR_{self \rightarrow other}$. This model generates the prediction that, holding all else constant, perceived increases in how another individual values the self (e.g., via benefit delivery) leads to an elevated WTR toward that individual and downstream motivations to cooperate with that individual.

1.2. Acknowledgement of benefit delivery: the function of gratitude

How might the increase in welfare valuation that occurs after another person has delivered a benefit be communicated to initiate or continue potential cooperative ventures? If benefit delivery is met with no response by the beneficiary, then benefactors are left uncertain about whether their action was perceived as holding any benefit, or for that matter, cost. A beneficiary's lack of response indicating benefit conferral (or cost imposition) translates into a lost opportunity to positively reinforce fitness-enhancing behaviors directed toward the self. All else equal, an adaptation that strategically operated to prolong the receipt of benefits would have gained an advantage over systems, for instance, that forewent such opportunities. We claim that the system(s) underlying gratitude is such an adaptation. The actions that gratitude motivates communicate to a benefactor the receipt of benefits and can strategically foreshadow the beneficiary's intent to return benefits. That is, expressions of gratitude effectively communicate, “I perceive your action to have benefited me, and, as a result, I have increased how I value you, and am likely to benefit you in the future should an opportunity arise.” The various intonations and expressions of gratitude allow for the communication of events ranging from extremely large benefit delivery (“I thank you with all of my heart”) all the way down to the sarcasm expressed when no benefit is realized, or even when a cost is imposed (“Gee, thanks”).

Gratitude, as an adaptation that positively reinforces benefit delivery, can be seen as antithetical to anger, an emotion that negatively reinforces cost imposition. Whereas anger—that is, the actions, communications, et cetera that anger causes—seems to function strategically as a recalibration mechanism when another person *devalues* one's welfare (Sell et al., 2009), gratitude seems to function strategically

when another person *values* one's welfare (see also Lim, 2012). Here, we test the proposal that gratitude is yoked to the WTR evaluation system and is activated upon detecting that another individual holds a higher WTR for the self than expected. As can be seen in Fig. 1, we predict that increases in gratitude should be coincident with increases in welfare evaluation.

1.3. The present experiment

The WTR-based model of cooperation discussed above suggests that when a stranger confers a benefit (that is, they reveal that they regard the self in a positive manner), the self's WTR value for the stranger increases, leading to feelings of gratitude and motivations to cooperate in future interactions. Here, we empirically evaluate this model. We employed a modified version of the well-established Cyberball game, a paradigm that manipulates social inclusion (Williams, Cheung, & Choi, 2000). We considered Cyberball a suitable task for our purpose here because we wanted to create a situation that elicits strong emotional reactions and to measure reactions toward both a benefactor and neutral (control) bystander. In accord with previous findings, we predicted that the intentional and costly delivery of benefits from a stranger toward a participant will predict the participant's downstream motivations to return benefits to the stranger, desires to befriend the stranger, and levels of gratitude reported toward the stranger.

With respect to valuation, we predicted that the intentional and costly delivery of benefits by a stranger will lead to increases in magnitudes of WTR, that is, positive Δ WTR. Whereas before the stranger's kind act the participant might have valued the stranger minimally, after the stranger's kind act, the participant should value the stranger more. That is, benefit delivery by a stranger should translate into an increase in the participant's WTR toward the stranger, $WTR_{\text{participant} \rightarrow \text{stranger}}$.

Further we tested two predictions regarding *how* the underlying welfare valuation system regulates the observed cooperative behavior and reported gratitude. As discussed above, current values of WTR are posited to be the critical regulatory variable predicting cooperative motivations. Thus, we predicted that follow-up WTR values—WTR values measured after benefit delivery—would predict DG allocations and desires for future interactions, not baseline WTR values or the extent to which WTR values changed. However, we predicted that levels of gratitude would correspond with *changes* in WTR values. That is, if gratitude's function is to communicate the magnitude of perceived

benefit received from another and the corresponding increasing in one's own WTR toward the benefactor, then reports of gratitude should be sensitive to changes in WTR.

2. Materials and methods

2.1. Participants

257 Introductory Psychology students (126 males; mean age 19.23; $SD = 2.48$) at the University of Miami participated in this study for course credit and \$10 compensation. 16 additional students did not complete the experiment due to computer network failures or because they knew a confederate. Their data were excluded from analyses.

2.2. Materials and procedure

2.2.1. Introductions

Upon arrival, each participant and three confederates waited together in a common area and were led to a room to get acquainted before beginning the experiment. Everyone was photographed, and these photographs were uploaded for use in the computer games. After introductions, the researcher led the participant and confederates to separate rooms to begin the computer games. (In fact, the three confederates, who were blind to condition and their role in the computer games, waited together in one room for the entire session.)

2.2.2. Baseline WTR

Participants first completed a measure of WTR for each of the three other players. The WTR measure (adapted from Jones & Rachlin, 2009) instructed participants to make 10 hypothetical decisions about whether they would prefer to receive a certain amount of money for themselves or to have \$75 go to the other player. Each decision proposed a different amount of money to the participant, starting at \$85, decreasing in \$10 increments, and ending at \$0. WTR is calculated by finding the indifference (or switch) point on the scale, defined as the average between the last amount the participant selected for him/herself and the first amount forgone to give \$75 to the other person, divided by \$75. This gives a possible WTR range of 0.00 to 1.13. We did not calculate a WTR score if a participant provided more than 1 switch-point over the course of the 10 decisions.

Efforts to develop and validate a measure designed to assess one's welfare trade-off ratio for another have shown WTR to have robust reliability and validity (Delton, 2010; Delton & Robertson, 2016). For instance, Delton (2010) varied the range and monetary value of the WTR scale and found test-retest reliability. Furthermore, friends and acquaintances accurately estimate the WTRs they hold for one another: Beliefs regarding how a friend values oneself closely matches the friend's actual reported WTR toward the self (Delton, 2010). Construct validity for the WTR has been demonstrated cross-culturally with both Argentinian college students (Delton, 2010) and the Shuar of Ecuador (Lim, 2012).

2.2.3. Cyberball

Next, participants played a modified version of Cyberball, a computerized game of catch in which a person can be included or excluded (Williams et al., 2000). The game began by assigning three players to be a Standard Player (SP) and one player to be Treasurer; the participant was always a SP. We refer to the other Standard Players as Standard Player #1 (SP1) and Standard Player #2 (SP2). Participants were told the rules of the game: the Treasurer earned \$0.50 for tossing the ball to any of the three SPs, while a SP earned \$0.50 only when tossing the ball to the Treasurer (not when tossing it to the other SPs). The incentive in the game was therefore to toss the ball to the Treasurer, and thus earn as much money as possible. The game continued until \$30 had been distributed (the Treasurer always earned \$15.00), though this termination criterion was not made known to participants.

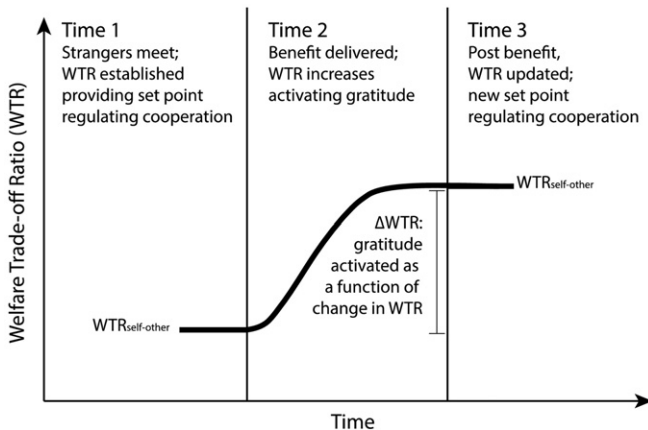


Fig. 1. A model of WTR-based cooperation. At Time 1, when two strangers meet, they each generate a WTR. Baseline WTR will depend on person specific variables of both the "self" and "other". At Time 2, when the other person has delivered a benefit to the self, WTR self-to-other increases, all things equal. The magnitude of the change in WTR reflects the change in perceptions of how the other individual values the self. New information regarding how that other individual values the self then updates the current value of WTR self-to-other, the new magnitude which then regulates cooperative versus competitive effort.

Participants were randomly assigned to one of three experimental conditions: Exclusion, Inclusion, or Exclusion then Inclusion. Participants in the Exclusion condition never received the ball for the entire game. Participants in the Inclusion condition were tossed the ball at equal odds as the other SPs. Participants in the Exclusion then Inclusion condition did not receive the ball for the first half of the game, that is, until \$15.00 had been distributed among the Treasurer, SP1, and SP2. After this point, participants were then included by one of the SPs (whom we term “Benefactor”). Although the Treasurer continued to exclude the participant, the Benefactor—and only the Benefactor—tossed the ball to the participant 80% of the time, thus enabling the participant to earn money by tossing the ball to the Treasurer (see Supplementary Information 1 in Appendix for condition pay-off structures).

2.2.4. Cyberball survey

Via self-report, we next assessed the degree to which participants felt included versus excluded during the first and second half of the Cyberball game. Specifically, using a 7-point Likert-type scale (1 = *strongly disagree*; 7 = *strongly agree*), we assessed the degree to which participants experienced the following 12 sentiments during the first and second half of the Cyberball game: *included, accepted, welcomed, noticed, considered, equally treated, excluded, rejected, isolated, ignored, not considered, and not equally treated* (presented in randomized order). We aggregated responses and used these data as manipulation checks. They show that our experimental conditions indeed varied as expected (see Supplementary Information 2 in Appendix). We also measured how included versus excluded participants felt by each of the other players in their condition (“I felt included by this player”; “I felt excluded by this player”), again rated on a 7-point Likert-type scale (1 = *strongly disagree*; 7 = *strongly agree*). Data showed that reactions to players varied in the expected ways (see Supplementary Information 3 in Appendix for analyses).

We also asked participants to indicate how grateful, thankful, appreciative, angry, and disgusted they felt toward or by each player, rated on the same 7-point agreement scale. Three additional questions were “I felt like I wanted to repay this player”, “I felt that I was obligated and had to repay this player”, and “I felt indebted (like I owed something) to this player”. We aggregated responses to questions about feeling “grateful,” “thankful,” and “appreciative” toward each other player into a measure termed Gratitude (alphas > 0.90 for Gratitude toward each player).

2.2.5. Follow-up WTR

Participants completed the same WTR measure for each player (see *Baseline WTR* section above).

2.2.6. Dictator Game

Next, all participants were led to believe that they had been randomly assigned to the role of Dictator in a multi-person Dictator Game that included the same three players from Cyberball. Participants were given \$10 and told they could keep the entire endowment or divide it among all four players in any manner. The three confederates were assigned to the role of recipients and had no influence on the outcome. Half of the participants were told their identity as Dictator would be known to the other participants (Identified condition) whereas the other half was told their identity would not be known. We included these separate conditions to determine whether decisions to allocate resources depended on whether their identity was known or not (Anonymous condition; see Supplementary Information 4 in Appendix).

2.2.7. Final survey

Last, participants were asked a variety of follow-up questions including which player (if any) they would prefer as a friend, and which player (if any) they would prefer to work with again as a partner if given the chance. Participants were then debriefed and dismissed.

2.3. Data analyses

With respect to terminology and labeling, in all conditions, we refer to Standard Players SP1 and SP2. In the Exclusion and Inclusion Cyberball conditions, these two players were programmed to exhibit similar behaviors. In the Exclusion-then-Inclusion Cyberball condition, one SP acts as a Benefactor, delivering benefits to the subject. We designated this SP as SP1; SP2 was programmed to behave similarly to the other SPs in the Exclusion and Inclusion Cyberball conditions.

2.3.1. WTR analyses

For each participant, we computed the difference between baseline WTR and follow-up WTR for each of the three other players (the Treasurer and 2 Standard Players). This difference is labeled as WTR change, or Δ WTR. When positive, Δ WTR indicates that the participant’s valuation of the other player’s welfare increased. When negative, Δ WTR indicates that the participant’s valuation of the other player’s welfare decreased. Change in WTR was calculated for 221 of 257 participants who provided follow-up WTR values for all players and a baseline WTR for at least one player (see Supplementary Information 5 in Appendix).

2.3.2. Gratitude analyses

Data from 161 subjects are used in analyses involving our gratitude measure (Exclusion then Inclusion: $N = 47$; Exclusion: $N = 73$; Inclusion: $N = 41$). Errors in the computer scripts for two conditions resulted in the failure to collect answers to follow-up questions regarding gratitude for 96 participants. For this reason, our N s vary across analyses. Computer games were programmed using E-Prime 2.0. Statistical analyses report two-tailed p -values.

3. Results

3.1. Benefit delivery by a stranger increases motivations to return benefits and befriend

We predicted that after the game of Cyberball, participants who had been assigned to the Exclusion then Inclusion condition would preferentially allocate benefits in a Dictator Game to the Benefactor over the other players. We conducted a repeated measures GLM analysis entering the amount of money allocated to Self, Treasurer, SP1, and SP2 as the within-subjects factor, and Cyberball condition as a between-subjects factor. We found a significant interaction between player type and Cyberball condition on amounts allocated in the Dictator Game, $F(6,458) = 22.20$, $p < 0.001$, partial $\eta^2 = 0.23$ (Fig. 2). Of interest here, in the Exclusion then Inclusion condition, participants allocated more money to SP1, the Benefactor, ($M = \$2.65$; $SD = \$1.53$) than to SP2 ($M = \1.01; $SD = \$1.10$; $t(74) = 7.46$, $p < 0.001$, Cohen’s $d = 1.23$), whereas in the Exclusion and Inclusion conditions, money allocated to SP1 and SP2 did not differ ($ps > 0.22$).

As can be seen by comparing allocations made in the Exclusion and Exclusion then Inclusion conditions, the increase in money allocated to SP1 (the Benefactor) in the Exclusion then Inclusion condition came at the cost of lower allocations to Self, rather than merely lower allocations to the other players. Indeed, whereas the amounts allocated to SP2 and the Treasurer did not differ between the Exclusion and Exclusion then Inclusion conditions (SP2: $t(152) = 0.94$, $p = 0.35$; Treasurer: $t(152) = 0.93$, $p = 0.35$), the amount allocated to Self ($t(152) = 2.82$, $p = 0.005$, Cohen’s $d = 0.45$) and SP1 did ($t(152) = 7.21$, $p < 0.001$, Cohen’s $d = 1.16$). Participants apparently used money from their own earnings to increase their allocations to their Benefactors. As discussed in Supplementary Information 4 (see Appendix), conditions of anonymity in the Dictator Game increased the amount allocated to Self, but reduced the amount allocated only to Treasurers and SPs who did not act as Benefactors, not to SPs who were Benefactors (see Supplementary Information Fig. S11).

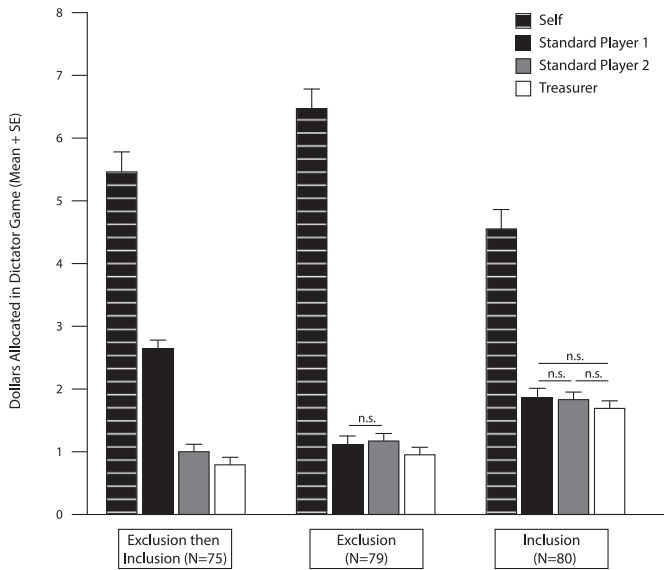


Fig. 2. Allocations made in the Dictator Game to players across Cyberball conditions. Note: Standard Player 1 is the Benefactor in the Exclusion then Inclusion condition.

At the end of the experiment, participants were asked to indicate which player they would select as a friend, if any. In the Exclusion then Inclusion condition, of the participants who selected a player, a significantly greater percentage (85.7%) selected SP1 (the Benefactor) than expected by chance alone whereas selection for SP1 did not deviate from chance in the Exclusion (38.9%) or the Inclusion (29.2%) conditions. A similar pattern held for responses to the questions of whom the participant would prefer to work with again as a partner (Supplementary Information 6 in Appendix for chi-square analyses and Fig. S12).

Our results replicate past findings and show that benefit delivery leads to downstream preferential treatment of a benefactor. Our next line of inquiry relates to the role of welfare valuation, specifically whether, as predicted, benefit delivery influences magnitudes of WTR, and whether WTRs, in turn, predict benefit return.

3.2. Benefit delivery between strangers coincides with increases in WTR

For subjects in the Exclusion then Inclusion condition, we expected that after being excluded in the Cyberball game, the costly and intentional inclusion by the Benefactor would lead to a greater increase in participants' WTR for the Benefactor as compared to the change in WTR for the other players in the Exclusion then Inclusion condition and the players in the Inclusion and Exclusion conditions. We conducted a repeated measures GLM analysis with Δ WTR for the Treasurer, SP1, and SP2 as the within-subjects factor and Cyberball experimental condition as the between-subjects factor. Multivariate tests indicated a significant interaction between player type and experimental condition, $F(4,434) = 27.27, p < 0.001$, partial $\eta^2 = 0.20$ (see Fig. 3). This interaction remained significant even after controlling for amount earned in the Cyberball game, $F(4,432) = 22.38, p < 0.001$, partial $\eta^2 = 0.17$.

Narrowing our analysis to just the Exclusion then Inclusion condition, we conducted a repeated measures GLM analysis entering Δ WTR for the three players as the within-subjects factor. Multivariate tests revealed a main effect of player type, $F(2,72) = 49.51, p < 0.001$, partial $\eta^2 = 0.58$. Paired samples *t*-tests revealed that participants' Δ WTR for the Benefactor (SP1) ($M = 0.08; SD = 0.30, N = 74$) was significantly greater than participants' Δ WTR for SP2 ($M = -0.27; SD = 0.30; t(73) = 7.85, p < 0.001$, Cohen's $d = 1.17$) and for the Treasurer ($M = -0.33; SD = 0.31; t(73) = 9.99, p < 0.001$, Cohen's $d = 1.34$). Furthermore, participants' Δ WTR for the Treasurer was significantly greater—that is, even

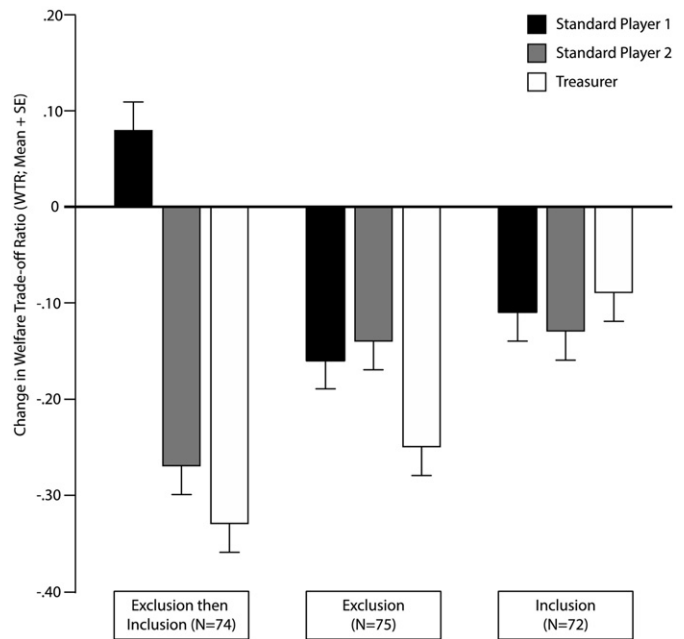


Fig. 3. Change in welfare trade-off ratio (WTR) toward each player across Cyberball conditions. The only increase in WTR occurred toward SP1 (the Benefactor) in the Exclusion then Inclusion condition. WTRs toward all other players decreased. There was no statistical difference in WTR change values between any of the Standard Players within or between the Inclusion and Exclusion conditions. Note: in the Exclusion then Inclusion condition, Standard Player 1 is the Benefactor.

more negative—than participants' Δ WTR for SP2, $t(73) = 2.58, p = 0.012$, Cohen's $d = 0.20$. All changes in WTR were significantly different from zero (see Table S13 in Information 5 in Appendix).

Differences in WTR changes toward the different players cannot be explained by differences in baseline WTR. A repeated measures GLM entering baseline WTR for the Treasurer, SP1, and SP2 as a within-subject factor and Cyberball condition as a between-subjects factor revealed no significant difference in participants' baseline WTRs for any of the players across conditions, $F(4,434) = 0.57, p = 0.682; N = 221$. Mean baseline WTR across 3 players in 3 conditions ranged from 0.50 to 0.57 (all SEs = 0.04).

As one might suspect, then, differences in the magnitude of WTR change are a result of the differences in follow-up WTR magnitudes. We replicated the above analyses, conducting a repeated measures GLM analysis with follow-up WTR for the Treasurer, SP1, and SP2 as the within-subjects factor and Cyberball experimental condition as the between-subjects factor, and found a significant interaction between player type and experimental condition, $F(4,438) = 35.83, p < 0.001$, partial $\eta^2 = 0.25$ (see Fig. 4).

Taken together, these results show that benefit delivery influences both follow-up WTR estimates and the magnitude of change from baseline to follow-up. This generates the question of how these two variables relate to motivations to deliver benefits in the Dictator Game.

3.3. Does follow-up magnitude of WTR and/or the extent to which WTR changes predict downstream motivations to return benefits?

Above we demonstrated that the costly and intentional delivery of benefits by a stranger coincides with an increase in the size of benefit returned to that stranger when playing the Dictator Game and an increase in WTR for the stranger. One of our goals here is to better understand how a system that assesses value regulates reciprocity. We predicted that a participant's newly updated WTR toward the stranger, which results from the stranger's behavior during Cyberball, should predict decisions to cooperate downstream. This is what we found. Follow-up WTR values significantly predicted the amount allocated to Standard

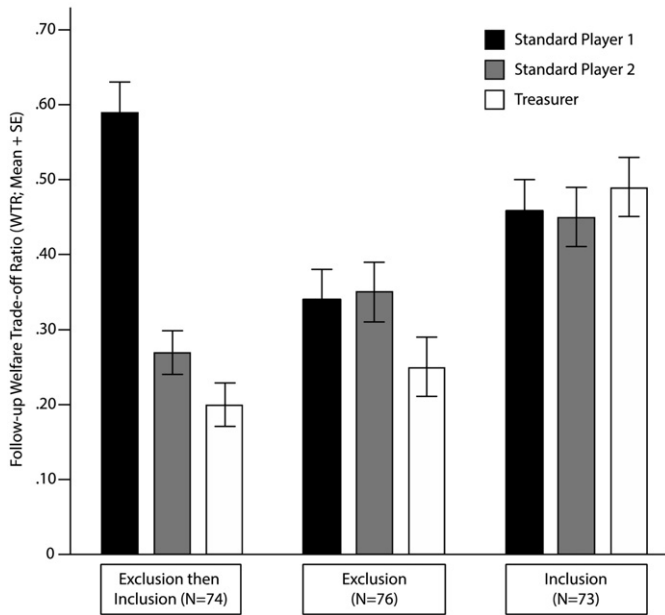


Fig. 4. Follow-up welfare trade-off ratio (WTR) toward each player across Cyberball conditions. In the Exclusion then Inclusion condition, follow-up WTR was greatest for the Benefactor (SP1), followed by SP2, and then the treasurer. In the Exclusion condition, follow-up WTR was lower for the treasurer as compared to either of the two SPs. In the Inclusion condition, follow-up WTRs did not differ between the players. Note: in the Exclusion then Inclusion condition, Standard Player 1 is the Benefactor.

Player 1 across all conditions, $r = 0.48$, $p < 0.001$, $N = 215$, and in each condition separately, Exclude-then-Include (in which SP1 was the Benefactor): $r = 0.34$, $p = 0.004$, $N = 69$; Include: $r = 0.43$, $p < 0.001$, $N = 72$; Exclude: $r = 0.54$, $p < 0.001$, $N = 74$. Similarly, collapsing across conditions, the Follow-up WTRs toward SP2 and the Treasurer predicted allocations in the Dictator Game: SP2: $r = 0.52$, $p < 0.001$, $N = 218$; Treasurer: $r = 0.53$, $p < 0.001$, $N = 218$.

However, across conditions, Δ WTR—the degree of change in the welfare trade-off ratio—also correlated with amount participants allocated toward SP1: $r = 0.31$, $p < 0.001$, $N = 205$, SP2: $r = 0.23$, $p = 0.001$, $N = 205$; and Treasurers: $r = 0.22$, $p = 0.002$, $N = 205$, raising the question of whether the relative change in WTR or the magnitude of the updated WTR toward another person best predicts cooperative behavior. Controlling for Δ WTR, follow-up WTR continues to predict allocations in the Dictator Game, SP1: *partial* $r = 0.43$, $p < 0.001$, $N = 205$; SP2: *partial* $r = 0.47$, $p < 0.001$, $N = 205$; Treasurer: *partial* $r = 0.49$, $p < 0.001$, $N = 205$. However, controlling for follow-up WTR, Δ WTR does not, SP1: *partial* $r = 0.09$, $p = 0.192$, SP2: *partial* $r = 0.03$, $p = 0.639$; Treasurer: *partial* $r = -0.03$, $p = 0.659$. Thus, the updated WTR magnitude—how one values another person right now—is the variable that best predicts cooperative behavior.

To this point, we have shown that (i) benefit delivery by a benefactor predicts a beneficiary's return of benefits to the benefactor; (ii) benefit delivery by a benefactor positively changes the WTR value the beneficiary holds for the benefactor, leading to an increased follow-up WTR magnitude; and (iii) the resulting WTR magnitude, not the amount by which WTR changes, predicts the allocation of money in the Dictator Game. However, there is a large body of research showing that benefit delivery also leads to gratitude and that feelings of gratitude are coincident with benefit return. How does gratitude fit into the picture of a WTR-based system of cooperation? We first replicate previous findings showing that benefit delivery generates gratitude and that gratitude predicts downstream benefit return. We then evaluate a novel prediction regarding how gratitude relates to WTR, and finally, using structural equation modeling, we evaluate the unique roles gratitude and WTR play in motivations to deliver benefits.

3.4. Benefit delivery by a stranger generates gratitude

We predicted that in the Exclusion then Inclusion condition, the costly and intentional delivery of benefits by SP1 (the Benefactor) would lead to a greater increase in participants' gratitude for the Benefactor as compared to the gratitude for the other players in this condition and the players in the Inclusion and Exclusion conditions. We conducted a repeated measures GLM analysis entering gratitude for the Treasurer, SP1, and SP2 as a within-subjects factor and Cyberball condition as the between-subjects factor. Multivariate tests indicated a significant interaction between player type and experimental condition, $F(4,314) = 154.93$, $p < 0.001$, *partial* $\eta^2 = 0.66$ (see Fig. 5).

Narrowing our analysis to just the Exclusion then Inclusion condition, we conducted a repeated measures GLM analysis entering gratitude for the three players as the within-subjects factor. Multivariate tests revealed a main effect of player type, $F(2,45) = 254.20$, $p < 0.001$, *partial* $\eta^2 = 0.92$, $N = 47$. Participants reported significantly greater gratitude toward the Benefactor ($M = 6.16$; $SD = 0.97$) than both SP2 ($M = 1.75$; $SD = 0.93$; $t(46) = 19.93$, $p < 0.001$; Cohen's $d = 4.64$) and the Treasurer ($M = 1.60$; $SD = 0.77$; $t(46) = 22.18$, $p < 0.001$; Cohen's $d = 5.21$). There was no difference between gratitude toward SP2 and the Treasurer ($p = 0.23$). See Supplementary Information 7 in Appendix for additional analyses of other conditions.

3.5. Gratitude predicts allocations in the Dictator Game

Across conditions, the amount of gratitude subjects felt toward each player predicted the amount of money allocated to that player in the Dictator Game: Standard Player 1 (note: this group includes the Benefactors in the Exclude-then-Include condition), $r = 0.48$, $p < 0.001$, $N = 145$; Standard Player 2, $r = 0.26$, $p = 0.002$, $N = 145$; Treasurer, $r = 0.47$, $p < 0.001$, $N = 234$. This finding replicates previous studies showing that gratitude is associated with greater cooperation. However, in the model we propose, we are suggesting that updated WTR estimates are the variables that regulate cooperative motivations, not gratitude.

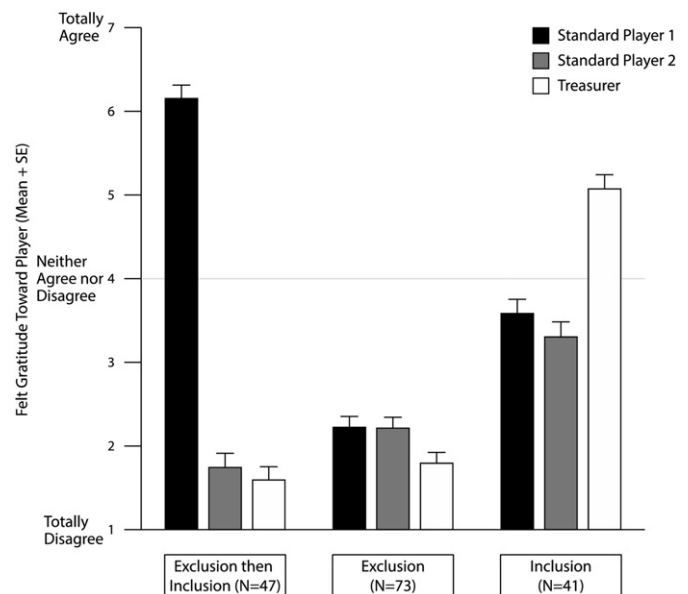


Fig. 5. Reported gratitude toward each player across Cyberball conditions. The two players toward whom participants felt grateful were the Benefactor in the Exclusion then Inclusion condition ($M = 6.16$; $SE = 0.16$) and the Treasurer in the Inclusion condition ($M = 5.08$; $SE = 0.17$).

3.6. Does receiving help from a Benefactor increase cooperation in the Dictator Game via the intermediate effects on gratitude, WTR, or both?

We analyzed a mediation model with two mediators, gratitude and follow-up WTR, and a multi-categorical predictor (Cyberball condition) using the 'lavaan' package in R 3.3.1 (Rosseel, 2012). We used the Exclusion-then-Inclusion as our reference/baseline condition. Cyberball condition was dummy coded with two vectors: one indicating the effect of the Exclusion condition relative to the Exclusion-then-Inclusion condition; and another indicating the effect of the Inclusion condition relative to the Exclusion-then-Inclusion condition. We only used data for SP1 allowing us to compare outcomes relating to the Benefactors (SP1) in the Exclusion-then-Inclusion condition to outcomes of the SP1 in the Exclusion condition and outcomes of the SP1 in the Inclusion condition. Missing data were handled using full information maximum likelihood (FIML) and confidence intervals were estimated with 1000 bootstrap samples (Hayes & Preacher, 2014; Preacher & Hayes, 2008). The model was used to estimate the direct effects of gratitude and WTR on dictator game behavior. The path from gratitude to dictator game behavior is the effect of gratitude on DG behavior when all other conditions are 0, thus revealing the effect of the baseline condition, which in this model is the experimental Exclude-then-Include condition. (We note for interpretation and future research design that the WTR scale consists of hypothetical monetary decisions, and our measure of cooperation was amount of money given in a real dictator game, opening the door to effects of method covariance.)

As shown in Fig. 6, the direct effect of follow-up WTR on dictator game behavior was significant, $b = 1.456, SE = 0.24, z = 6.061, p < 0.001, 95\% CI [0.990, 1.912]$, indicating an expected increase in approximately \$1.46 when going from $WTR = 0$ to $WTR = 1$ in the Exclusion-then-Inclusion baseline condition (note: this represents nearly the entire range of the WTR scale). By contrast, the direct effect of gratitude on dictator game behavior was not significant, $b = 0.057, SE = 0.077, z = 0.735, p = 0.462, 95\% CI [-0.098, 0.213]$ indicating that, for the Exclusion-then-Inclusion (baseline) condition, when controlling for follow-up WTR, gratitude was not predictive of dictator game giving. Further, the indirect effects from Exclusion, $b = -0.395, SE = 0.096, z = -3.745, p < 0.001, 95\% CI [-0.555, -0.195]$, and Inclusion, $b = -0.228, SE = 0.088, z = -2.586, p = 0.010, 95\% CI [-0.428, -0.071]$, predicted dictator game giving via follow-up WTR, but not via gratitude (for remaining direct effects and indirect effects, see Supplementary Information #8 in Appendix).

3.7. If gratitude doesn't predict cooperation, what does gratitude do?

In our model, we predicted that change in WTR should predict levels of gratitude. That is, the perception that someone values you more than expected, leading you to increase your WTR toward them, generates gratitude. When we examine participants in the Exclusion-then-Inclusion condition and their reactions toward Benefactors (SP1), we find a significant correlation between ΔWTR and gratitude, $r = 0.55, p < 0.001, N = 42$. However, there is also a significant correlation between follow-up WTR and gratitude, $r = 0.43, p = 0.005, N = 42$. If, as we propose, gratitude functions to reinforce benefit delivery and to communicate to a benefactor the receipt of larger-than-expected benefits and an increase in one's own WTR, then change in WTR should predict gratitude. However, if gratitude functions to, for instance, communicate one's current magnitude of valuation for another, then follow-up WTR should best predict gratitude. Controlling for follow-up WTR, ΔWTR continues to predict gratitude, partial $r = 0.42, p = 0.006, N = 42$. By contrast, controlling for ΔWTR , follow-up WTR no longer predicts gratitude, partial $r = 0.20, p = 0.213, N = 42$. Thus, ΔWTR appears to be the underlying representation contributing to gratitude.

4. Discussion

Here, we tested an information-processing model of the proximate regulation of cooperative relationships. We found evidence in support of three main predictions: (i) benefit delivery by a stranger leads to an increase in welfare valuation; (ii) updated magnitudes of welfare valuation predict downstream cooperative behavior; and (iii) feelings of gratitude are coincident with positive changes in welfare valuation. This model of how cooperative relationships are established, depicted in Fig. 1, hinges on estimated changes in welfare valuation. Events that confer benefits are translated into internal representations of welfare valuation and compared against stored WTR values. Situations that lead to positive changes in WTR values can strategically activate motivations to select benefactors as social partners. Conversely, negative changes in WTR values can strategically activate motivations to avoid particular individuals as social partners or to recalibrate upwards another's WTR for oneself (see Sell et al., 2009).

Of potential interest, the only individual for whom participants experienced an increase in WTR was the Benefactor in the Exclusion-then-Inclusion condition (see Fig. 3). Changes in WTRs toward the other Standard Player and the Treasurer in this condition were all

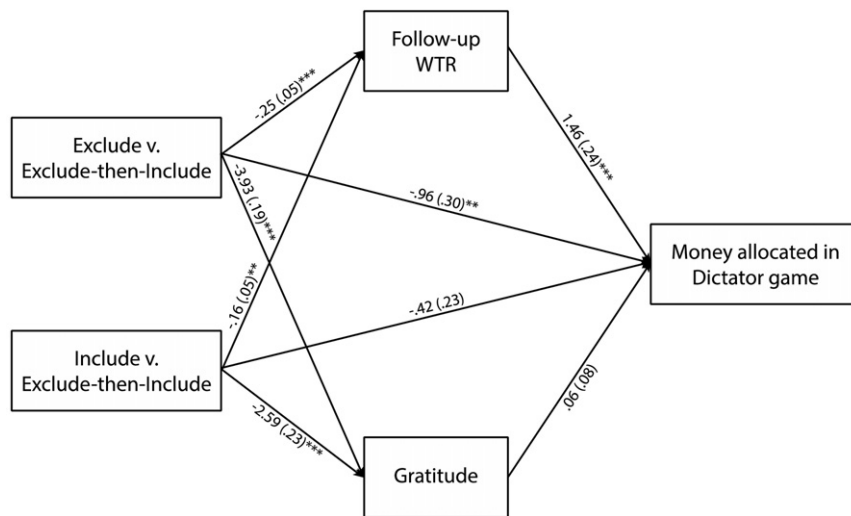


Fig. 6. A mediation model. The Exclude-then-Include condition is the reference/baseline condition; other conditions are dummy coded. Values are standardized B (standard error). Model shows the effect of follow-up WTR on dictator game behavior controlling for gratitude and effect of gratitude on dictator game behavior controlling for follow-up WTR. Paths indicate values for Exclude-then-Include baseline condition (i.e., when other conditions are 0). Model also shows direct effects of Exclude and Include condition relative to the Exclude-then-Include condition on follow-up WTR, gratitude, and dictator game behavior (see Supplementary Information 8, which also includes a discussion of indirect and total effects).

negative, an outcome that might have been attributed to the fact that there was a contrast effect in this condition—one person was generous; the others were not. But, changes in WTRs in the other conditions were also negative. This isn't surprising for the Exclude condition, but even the Include condition revealed negative changes in WTR, a condition in which the subject benefitted like the other standard players. Though certainly requiring additional investigation, one possibility is that initial WTRs take on diffuse, uncertain values until an event occurs that reduces the uncertainty of how another individual values the self. It appears—at least from our data—that humans are a bit optimistic and biased toward cooperation as some have proposed (e.g., Delton et al., 2011). Baseline WTR values begin higher than those ultimately revealed by the standard players and treasurer in our Cyberball Inclusion condition. That is, at least in our payoff structure, baseline WTRs started higher than the value revealed by an individual acting to increase their own monetary gain. The standard players in the Inclusion condition, by throwing the ball back to the Treasurer for monetary gain, were providing information regarding self versus other valuation; the other player valued him/herself over the subject. The absence of any benefit delivering interaction decreased estimated values of WTR toward the others in the Inclusion condition (and other conditions as well). That is, WTRs toward other players declined over time as opportunities to deliver benefits occurred yet were intentionally not taken.

4.1. WTR-based reciprocity

The model of cooperation we tested is based on stored estimates of WTR, a continually updated summary variable computed from information regarding the self-directed costs and benefits of another individual's actions, kinship, and association value. Critically, in this model, there is no required long-term, high-fidelity storage of interaction history, only a summary variable, WTR, that is continually updated based on present interactions (and exposure to relevant kinship and other fitness-relevant cues). Once a WTR is updated due to an act of another, the act itself can be forgotten, suggesting fewer cognitive requirements in terms of memory storage and access. The classic reciprocity examples of blood-sharing in bats (Wilkinson, 1984) and food sharing and grooming in primates (de Waal, 1989; Gomes, Mundry, & Boesch, 2009; Seyfarth & Cheney, 1984), rather than resulting from stored histories of multiple interactions and information regarding temporal contingency, could instead result from a slightly more simple system that updates a single representation of value as discussed herein.

The fact that these species—and others—already have kin detection systems in place to guide kin-directed altruism (Silk, 2005) suggests they also have systems for assessing inter-individual value. Indeed, a single system that can handle kinship estimates of zero and higher is consistent with cooperative behavior observed between kin and non-kin in non-human primates (Schino & Aureli, 2009). This leads to the intriguing possibility that despite the different selective pressures giving rise to kin-based helping and helping due to mutual benefit, the computational architecture that implements them might be encompassed within the WTR-based system we outline here.

An important feature of the WTR-based system is that it can protect against exploitation. Increases and decreases in value for another serve to protect against incurring repeated costs when attempting to initiate a cooperative relationship. That is, a value-based representation can offer an explanation for how repeated exploitation is avoided. For instance, after A delivers a benefit to B, the absence of B's delivery of benefits back to A should lead, over time and depending on circumstance, to a decrease in A's WTR for B, lowering the probability that A will deliver future benefits to B.

4.2. Gratitude

This model points to the importance of gratitude as a key emotion in the initiation and maintenance of mutual-benefit cooperation in

humans. Gratitude, we suggest, serves an important function in acknowledging the receipt of (greater-than-expected) benefits and announcing that the beneficiary's regard for the benefactor's welfare has increased as a result. In this way, gratitude functions in a manner that is nearly opposite that of anger: Whereas expressed anger indicates that another individual has not valued one's welfare to the degree expected (see Sell et al., 2009), expressed gratitude indicates that another individual has valued one's welfare at a higher-than-expected level (Lim, 2012). Indeed, as compared to the correlation reported above between WTR change for the Benefactor and gratitude toward Benefactor among participants in the Exclusion then Inclusion condition ($r = 0.55$), the correlation between WTR change and anger was of nearly identical magnitude, but in the opposite direction: $r = -0.53$, $p < 0.001$, $N = 42$. In addition to playing a role in initiating relationships via delivery of benefits, gratitude might also serve a maintenance function in established relationships, reinforcing levels of current value and further reducing uncertainty. Indeed, recent evidence suggests this is the case (e.g., Algoe & Zhaoyang, 2016).

5. Conclusions

The development of information-processing models of cooperation can help to identify potential parameters that natural selection might have identified and refined to adaptively guide social behavior. There are likely alternate cognitive architectures that could, in principle, achieve the same outcomes as discussed herein. Here we proposed and tested a model that focused on WTR estimates and how changes in these estimates can lead to the formation of cooperative relationships and production of gratitude among non-kin in humans. The cognitive systems in place in many social species, such as those governing kin-based altruism, point to the intriguing possibility that WTR values might also explain patterns of reciprocity in non-humans as well.

Acknowledgments

This research was supported by a grant to D.L. and M.E.M. (subagreement no. 00008036) from UC Berkeley's Greater Good Science Center, which received funding from the John Templeton Foundation (Grant ID No. 15627, Berkeley award ref. #201121067-01) and by a separate grant to M.E.M. from the John Templeton Foundation (grant no. 29615).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.evolhumbehav.2017.08.003>.

References

- Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: The 'other-praising' emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology, 4*, 105–127.
- Algoe, S. B., & Zhaoyang, R. (2016). Positive psychology in context: Effects of expressing gratitude in ongoing relationships depend on perceptions of enactor responsiveness. *The Journal of Positive Psychology, 11*, 399–415.
- Bartlett, M. Y., Condon, P., Cruz, J., Baumann, J., & DeSteno, D. (2012). Gratitude: Prompting behaviors that build relationships. *Cognition & Emotion, 26*, 2–13.
- Bartlett, M. Y., & DeSteno, D. (2006). Gratitude and prosocial behavior: Helping when it costs you. *Psychological Science, 17*, 319–325.
- Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology, 13*, 171–195.
- Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 584–627). New York: Wiley.
- Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology, 64*, 201–229.
- de Waal, F. B. M. (1989). Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution, 18*, 433–459.
- Delton, A. W. (2010). *A psychological calculus for welfare tradeoffs*. (Doctoral dissertation) Santa Barbara: Department of Psychology, University of California.

- Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Science*, *108*, 13335–13340.
- Delton, A. W., & Robertson, T. E. (2016). How the mind makes welfare tradeoffs: Evolution, computation, and emotion. *Current Opinion in Psychology*, *7*, 12–16.
- DeSteno, D., Bartlett, M. Y., Baumann, J., Williams, L. A., & Dickens, L. (2010). Gratitude as moral sentiment: Emotion-guided cooperation in economic exchange. *Emotion*, *10*, 289–293.
- Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: An experimental investigation of gratitude and subjective well-being in daily life. *Journal of Personality and Social Psychology*, *84*, 377–389.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785–791.
- Forster, D. E., Pedersen, E. J., Smith, A., McCullough, M. E., & Lieberman, D. (2017). Benefit valuation predicts gratitude. *Evolution and Human Behavior*, *38*(1), 18–26.
- Gergen, K. J., Ellsworth, P., Maslach, C., & Seipel, M. (1975). Obligation, donor resources, and reactions to aid in three cultures. *Journal of Personality and Social Psychology*, *31*, 390–400.
- Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*, 153–172.
- Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proceedings of the Royal Society B*, *276*, 699–706.
- Griffin, A. S., & West, S. A. (2003). Kin discrimination and the benefit of helping in cooperatively breeding vertebrates. *Science*, *302*, 634–636.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. I, II. *Journal of Theoretical Biology*, *7*, 1–52.
- Hammerstein, P. (2003). *The genetic and cultural evolution of cooperation*. Cambridge: MIT Press.
- Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470.
- Jones, B. A., & Rachlin, H. (2009). Delay, probability, and social discounting in a public goods game. *Journal of Experimental Analysis of Behavior*, *91*, 61–73.
- Krasnow, M. M., Cosmides, L., Pedersen, E. J., & Tooby, J. (2012). What are punishment and reputation for? *PLoS ONE*, *7*, e45662.
- Krasnow, M. M., Delton, A. W., Tooby, J., & Cosmides, L. (2013). Meeting now suggests we will meet again: Implications for debates on the evolution of cooperation. *Scientific Reports*, *3*, 1747.
- Lieberman, D., Tooby, J., & Cosmides, L. (2007). The architecture of human kin detection. *Nature*, *445*, 727–731.
- Lim, J. (2012). *Welfare tradeoff ratios and emotions: Psychological foundations of human reciprocity*. (Doctoral dissertation) Santa Barbara: Department of Anthropology, University of California.
- McCullough, M. E., Kilpatrick, S. D., Emmons, R. A., & Larson, D. B. (2001). Is gratitude a moral affect? *Psychological Bulletin*, *127*, 249–266.
- McCullough, M. E., Pedersen, E. J., Tabak, B. A., & Carter, E. C. (2014). Conciliatory gestures promote forgiveness and reduce anger in humans. *Proceedings of the National Academy of Science*, *111*, 11211–11216.
- McCullough, M. E., Tsang, J., & Emmons, R. A. (2004). Gratitude in intermediate affective terrain: Links of grateful moods to individual differences and daily emotional experience. *Journal of Personality and Social Psychology*, *86*, 295–309.
- Nowak, M. A., & Roch, S. (2007). Upstream reciprocity and the evolution of gratitude. *Proceedings of the Royal Society of London B: Biological Sciences*, *274*, 605–610.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*, 1291–1298.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.
- Rand, D. G., Ohtsuki, H., & Nowak, M. A. (2009). Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *Journal of Theoretical Biology*, *256*, 45–57.
- Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, *70*, 901–908.
- Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.
- Schino, G., & Aureli, F. (2009). The relative roles of kinship and reciprocity in explaining primate altruism. *Ecology Letters*, *12*, 1–6.
- Schino, G., & Aureli, F. (2010). Primate reciprocity and its cognitive requirements. *Evolutionary Anthropology*, *19*, 130–135.
- Sell, A. (2011). The recalibrational theory and violent anger. *Aggression and Violent Behavior*, *16*, 381–389.
- Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Science*, *106*, 15073–15078.
- Seyfarth, R. M., & Cheney, D. L. (1984). Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature*, *308*, 541–543.
- Silk, J. B. (2005). The evolution of cooperation in primate groups. In H. Gintis, S. Bowles, R. Boyd, & E. Fehr (Eds.), *Moral sentiments and material interests: On the foundations of cooperation in economic life* (pp. 43–73). Cambridge: MIT Press.
- Sussman, R. W., & Cloninger, C. R. (2011). *Origins of altruism and cooperation*. New York: Springer.
- Sznycer, D. (2010). *Cognitive adaptations for calibrating welfare tradeoff motivations, with special reference to the emotion of shame*. (Doctoral dissertation) Santa Barbara: Department of Anthropology, University of California.
- Tesser, A., Gatewood, R., & Driver, M. (1968). Some determinants of gratitude. *Journal of Personality and Social Psychology*, *9*, 233–236.
- Tooby, J., & Cosmides, L. (1996). Friendship and the Banker's Paradox: Other pathways to the evolution of adaptations for altruism. In W. G. Runciman, J. M. Smith, & D. RIM (Eds.), *Evolution of social behaviour patterns in primates and man. Proceedings of the British Academy*, vol. 88. (pp. 119–143). London: British Academy.
- Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotion* (pp. 114–137) (3rd ed.). New York: Guilford.
- Tooby, J., Cosmides, L., Sell, A. N., Lieberman, D., & Sznycer, D. (2008). Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. In A. J. Elliott (Ed.), *Handbook of approach and avoidance motivation* (pp. 251–271). Mahwah, NJ: Lawrence Erlbaum Associates.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.
- Tsang, J. A. (2006). The effects of helper intention on gratitude and indebtedness. *Motivation and Emotion*, *30*, 198–204.
- Tybur, J., & Lieberman, D. (2016). Human pathogen avoidance adaptations. *Current Opinion in Psychology*, *7*, 6–11.
- Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, *308*, 181–184.
- Williams, K. D., Cheung, C. K., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the internet. *Journal of Personality and Social Psychology*, *79*, 748–762.