

The Topography of Generosity: Asymmetric Evaluations of Prosocial Actions

Nadav Klein and Nicholas Epley
University of Chicago

Prosociality is considered a virtue. Those who care for others are admired, whereas those who care only for themselves are despised. For one's reputation, it pays to be nice. Does it pay to be even nicer? Four experiments assess reputational inferences across the entire range of prosocial outcomes in zero-sum interactions, from completely selfish to completely selfless actions. We observed consistent nonlinear evaluations: Participants evaluated selfish actions more negatively than equitable actions, but they did not evaluate selfless actions markedly more favorably than equitable actions. This asymptotic pattern reflected monotonic evaluations for increasingly selfish actions and insensitivity to increasingly selfless actions. It pays to be nice but not to be really nice. Additional experiments suggest that this pattern stems partly from failing to make spontaneous comparisons between varying degrees of selflessness. We suggest that these reputational incentives could guide social norms, encouraging equitable actions but discouraging extremely selfless actions.

Keywords: prosocial behavior, reputation, fairness, social judgment, attribution

Supplemental materials: <http://dx.doi.org/10.1037/xge0000025.supp>

It is difficult to think of generosity and selfishness without also thinking about praise and scorn, respectively. In philosophy, Descartes, Hume, and Aquinas all identified self-sacrifice for the benefit of others as a major ethical accomplishment in a person's life. In religion, the Bible exalts generosity as the most important spiritual virtue (Corinthians 1:13) and links selfishness with evil (James 3:16; Romans 2:8). Buddha went further, teaching that compassion (karuna) not merely is a virtue but is one of only two essential qualities for enlightenment (wisdom being the other). In literature, Ebenezer Scrooge becomes beloved only after exchanging his miserly ways for generosity toward others (Dickens, 1843). In modern life, no representative sample is needed to confirm that seemingly self-sacrificing figures like Mother Teresa and Mahatma Gandhi are admired whereas selfish actors like Jeff Skilling and Bernard Madoff are despised.

These examples reflect a widespread tendency to admire prosociality in others and to despise selfishness. A child who shares a bite of his cookie will be liked more than one who shares none of it. A wealthy person who gives 10% of her income to charity will

appear more caring than one who gives none of her income. A professor who offers half a workday to help students will be more beloved than a professor who offers none of the day. In terms of one's reputation, it pays to be nice.

Does it pay to be even nicer? That is, does behaving even more prosocially—benefiting others at an increasing cost to oneself—lead to an even more positive reputation? Is a child who gives away an entire cookie liked more than one who shares only a bite? Is the wealthy person who gives 50% of her income seen as more caring than one who gives away only 5%? Is a professor who gives an entire workday to help students admired more than one who gives only half a day?

These questions address a fundamental issue in social judgment: Does behaving more prosocially lead to a consistently more positive reputation across the entire range of prosocial actions from completely selfish to completely selfless? Existing research typically measures evaluations of only discrete points along this continuum, such as clearly selfish versus clearly generous actions (e.g., Almenberg, Dreber, Apicella, & Rand, 2011; Berman & Small, 2012; Gray, Ward, & Norton, 2014), without studying the entire range of possible behavior. These questions also address a fundamental issue for self-regulation: If a person wants to maintain a positive reputation, how exactly should he or she behave? Some acts of kindness could incur costs that are not repaid in reputational value. Knowing the reputational value of prosocial behavior may explain how people optimize the potential trade-off between reputational benefits and personal costs.

We conducted 11 experiments—eight described here and three described in the online supplemental materials—that address these questions. In each experiment, we study evaluations in zero-sum contexts where people can benefit either themselves or others in varying degrees. In these contexts, the precise degree of selfish versus selfless action is maximally clear.

This article was published Online First October 13, 2014.

Nadav Klein and Nicholas Epley, Booth School of Business, University of Chicago.

We thank the Booth School of Business for financial support; Shawn Fagan, Jill Jacobbi, Jasmine Kwong, Sean Alexander McWillie, Sarah Molouki, Megan Porter, Emily Shaw, and Whitney Westmorland for assistance in conducting the experiments; Daniel Gilbert and Thomas Gilovich for helpful comments on a draft of this article; and Rashida Black for allowing us to conduct an experiment during University of Chicago Orchestra concerts.

Correspondence concerning this article should be addressed to Nicholas Epley or Nadav Klein, Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637. E-mail: epley@chicagobooth.edu or nklein@chicagobooth.edu

Theoretical Possibilities

If prosociality is considered a virtue, then a person who shows more of it should earn a more positive reputation than a person who shows less. This predicts a relatively monotonic pattern across the range of prosocial actions, such that selfish actors are evaluated more negatively than moderately selfless (or equitable) actors, and extremely selfless actors (who maximally benefit others at a cost to the self) are evaluated more favorably than only moderately selfless actors (see Figure 1). Evaluations of a person's prosocial actions might therefore resemble judgments of other personal characteristics, such as height. As a person grows taller, the person is judged to be taller across the entire range of human height. At no point across the spectrum does a taller person stop looking taller or a shorter person stop looking shorter. On this account, it pays to be nice and pays more to be really nice.

As logical as this monotonic pattern may be, we believe there are at least two reasons to doubt that people's reputations actually conform to it. First, others' selfless behavior may become a threat to one's own sense of self-esteem, a threat that people may combat by derogating others' selfless actions (Fetchenhauer & Dunning, 2010; Minson & Monin, 2012; Monin, Sawyer, & Marquez, 2008). This predicts an inverted-*U* pattern, such that selfish actors are evaluated more negatively than equitable actors, but selfless actors who benefit others even more than themselves are evaluated more negatively than merely equitable actors (see Figure 1). On this account, it actually pays less to be really nice.

Second, to value outcomes of varying magnitudes rationally, these outcomes have to be evaluated in comparison to each other (Hsee & Zhang, 2010). People may indeed evaluate extremely prosocial actions more positively than moderately prosocial actions, but if they do not compare one against the other they may evaluate the two outcomes similarly. If evaluating a person who donated 5% of income to charity does not lead people to think of the possibility of giving away 50% to charity, then a mildly charitable person may be evaluated the same as an extremely charitable person when evaluated in isolation. Existing research demonstrates that value sensitivity depends partly on stimulus familiarity (Hsee & Zhang, 2010). The more familiar or knowledgeable a person is with a stimulus, the more sensitive eval-

uations will be to the objective value of the stimulus (Ariely & Loewenstein, 2000; Laming, 1997; Linville, 1982; Parducci, 1965; Stevens, 1975). For example, in one study the duration of a commute affected evaluations of that commute only for people who experienced this commute on a frequent basis (Morewedge, Kassam, Hsee, & Caruso, 2009). In another study, providing participants with distributional information about admissions test scores markedly increased discernment between low- and high-scoring candidates (Hsee, Loewenstein, Blount, & Bazerman, 1999). Likewise, norm theory suggests that when people evaluate familiar stimuli (versus unfamiliar stimuli), they are more likely to spontaneously generate comparison standards (Kahneman & Miller, 1986).

Variations of selfless behavior—in which people benefit others at varying costs to the self—are simply less common in everyday life than selfish behavior (Dawkins, 1976; Trivers, 1971). Americans, for instance, have donated roughly 2% of gross domestic product to charity consistently for the past 30 years while keeping the bulk of the remaining 98% for themselves (Giving USA, 2013). A recent review of dictator games ($N = 20,183$) found that more than 70% gave less than an even split, whereas only 13.1% gave more than an even split (Engel, 2011). Because selfless actions are relatively rare and unfamiliar, evaluations of selfless actions may be less sensitive to varying levels of generosity, leading them to be judged categorically rather than relatively. In contrast, selfish actions may be more likely to trigger spontaneous comparisons because they are both negative and common, implying that they would be evaluated in a broader perspective, in relation to the magnitude of selfishness. This predicts an asymptotic pattern: People judge increasing degrees of selfishness increasingly more negatively but do not judge increasing degrees of selflessness more positively (see Figure 1). On this account, it pays to be nice but pays no more to be really nice.

Overview of Experiments

We conducted a series of experiments to identify whether evaluations of prosocial actions follow a relatively monotonic, inverted-*U*, or asymptotic pattern. In each, participants evaluated a person who behaved relatively selfishly (benefiting only the self), selflessly (generously benefiting others at an increasing cost to oneself), or some point in between. Experiment 2 also tests whether people know how their prosocial actions are evaluated in the eyes of others, a crucial judgment for self-regulation. Experiments 4–6 test a possible explanation for the pattern of evaluations we observe.

A person's reputation may vary along many different dimensions, but existing research demonstrates that it typically varies along only two fundamental dimensions: warmth and competence (Fiske, Cuddy, & Glick, 2007; Judd, James-Hawkins, & Yzerbyt, 2005). Warmth is more dominant in judgment (Willis and Todorov, 2006; Wojciszke & Abele, 2008) and is related to other-oriented outcomes (e.g., friendliness, trustworthiness, morality), whereas competence is secondary and related to self-oriented outcomes (e.g., intelligence, talent, skill). We expected that prosocial actions would affect only reputations of warmth, but we measured competence in some of the studies in case the two were inversely related such that extreme acts of kindness are also seen as a sign of incompetence (Judd et al., 2005).

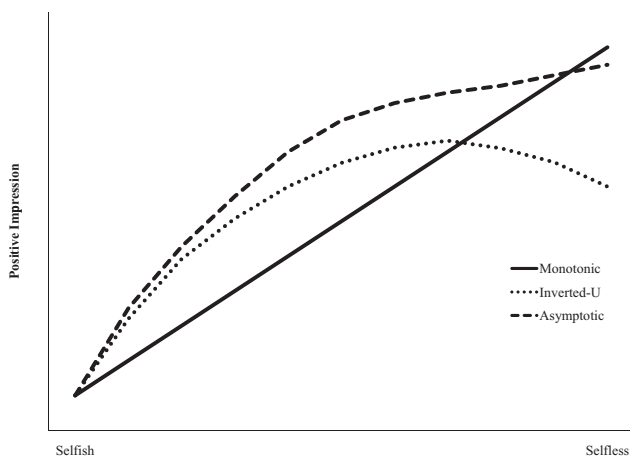


Figure 1. Theoretical relationships between prosociality and reputational inference.

Experiments 1a and 1b: Increasingly Generous Donations

We first tested our hypotheses in a situation in which people commonly witness and engage in a wide range of prosocial behavior. Specifically, our university orchestra concerts are free but include a suggested donation of \$10. Initial conversations with orchestra personnel verified that concertgoers vary in their behavior, with some donating nothing and some donating well above the suggested donation (in our sample, the highest donation was \$90 for a concert ticket). We asked actual concertgoers (Experiment 1a) or observers (Experiment 1b) to evaluate the character of a person who behaved increasingly generously (i.e., donated less, exactly, or more than the suggested donation in varying degrees).

Method

Experiment 1a. People who attended two university orchestra concerts received a survey in their program that they could complete voluntarily and return at the exit ($N = 102$, 57% female, $M_{\text{age}} = 47.7$ years). The orchestra suggests a \$10 donation at the concert, with donation baskets placed at the entrance. Thus, donating at concerts is optional.

The survey asked concertgoers to consider “Tom W.,” a typical classical music lover who attended a recent concert. The survey contained a generic description of Tom and then indicated that he donated \$0, \$10 (the suggested amount), \$20, or an unspecified amount. Using the fundamental dimensions identified by Fiske, Cuddy, Glick, and Xu (2002), participants then rated Tom’s character on five traits related to warmth (tolerant, warm, good-natured, sincere, and caring) and five traits related to competence (competent, confident, independent, competitive, and intelligent), in counterbalanced order (1 = *not at all*; 7 = *very much*). Participants then reported how much they personally donated at the concert. In the second concert only, participants ($n = 57$) estimated Tom’s annual salary at the survey’s end, to assess an alternative hypothesis that occurred to us only after the first concert (i.e., that Tom’s donation suggests something about his financial ability to donate rather than something about his character).

Experiment 1b. Participants ($N = 447$) from Amazon.com’s Mechanical Turk (MTurk) read about a situation that mirrored Experiment 1a, describing an orchestra that suggests a donation for admission to concerts. The prompt also indicated that “some people donate the suggested amount, some donate more, and some donate less.” We then manipulated the orchestra’s suggested donation to be no specified amount, \$10, \$20, or \$50. Participants then read the same description of Tom W. from Experiment 1a and that he donated \$0, \$5, \$10, \$20, or \$50 at the concert. Experiment 1b therefore varied not only the target’s behavior but also the expected prosocial conduct.

Results

Experiment 1a.

Past concert attendance. Participants attended 2.51 concerts on average during the academic year, and 57.2% of participants attended more than one concert.

Donation behavior. Ninety of our participants reported their own donations. A substantial portion (41.1%) donated the sug-

gested \$10. Among the rest, 48.9% donated less than \$10 (26.7% donated nothing), and 10.0% of participants donated more than \$10 (range = \$15–\$90). Thus, selfish behavior (donating less than the suggested amount) was approximately five times more common than generous behavior (donating more than the suggested amount) in this context. Consistent with the more general pattern of behavior in everyday life we cited in the introduction, relatively generous actions were more rare than relatively selfish actions.

Impressions. To test our main hypotheses, we first created two composites, one of the warmth items ($\alpha = .76$) and one of the competence items ($\alpha = .81$). Nine participants failed to answer at least one of the impression items (5 related to warmth, 6 related to competence). We did not include these participants with missing data when creating our composites. However, including them by simply averaging across the other items does not affect any of the following results in any meaningful way.

Competence ratings did not vary across conditions ($ts < 1$). Generosity was not seen as a sign of incompetence.

As shown in Table 1, Tom was judged more negatively when he behaved selfishly (donating \$0; $M = 3.51$, $SD = 1.31$) than when he behaved fairly (donating \$10; $M = 5.29$, $SD = 0.87$), $t(48) = 5.24$, $p < .001$, $d = 1.60$, but was not judged more favorably when he behaved generously (donating \$20; $M = 5.17$, $SD = 1.10$) than when he behaved merely fairly, $t(40) = 0.39$, $p = .70$, $d = 0.12$. This suggests an asymptotic pattern in evaluations of prosociality, whereby selfish actions are evaluated more negatively than fair actions, but generous actions are not evaluated more positively than fair actions. Interestingly, Tom’s reputation in the no information condition did not differ from the fair or generous conditions ($ts < 1.20$, $ps > .23$). This result likely occurred because participants simply assumed he donated the suggested amount (a logical inference given that 41.1% of participants actually donated the suggested amount).

Salary predictions. Predictions of Tom W.’s salary did not vary across conditions ($F < 1$). Participants inferred something about Tom W.’s character from his donation rather than about his financial ability to donate.¹

Experiment 1b. We again created a composite of the warmth items ($\alpha = .95$) and a composite of the competence items ($\alpha = .73$). Competence evaluations did not vary across suggested donation amounts, $F(3, 446) = 1.69$, $p = .17$, $\eta_p^2 = .01$. Larger donations led to more positive competence evaluations, $F(4, 446) = 8.51$, $p < .001$, $\eta_p^2 = .07$. The interaction was nonsignif-

¹ Given that participants in this experiment faced the same donation opportunity as the hypothetical Tom W. did, it is possible that participants’ own donation affected their evaluations. Our participants donated \$7.93, on average, and 67.0% of participants donated a positive amount. When Tom W. was described as donating \$20, most participants experienced an unfavorable social comparison and may therefore have failed to evaluate Tom W. more positively to maintain positive self-regard. However, we observed no significant correlation between participants’ own donation amounts and their evaluations of Tom W.’s warmth in the generous condition, $r(23) = -.09$, *ns*, suggesting no contrast effect in that condition. Instead, we observed equally negative correlations in both the fair, $r(19) = -.47$, $p = .07$, and the selfish conditions, $r(31) = -.44$, $p < .01$. This pattern suggests that participants’ own donation was affecting their evaluation of Tom W. but was doing so consistently in the selfish and fair conditions, making it an unlikely explanation for the difference in evaluations between these two conditions.

Table 1
Evaluations of Warmth and Competence by Donation Amount in Experiment 1a

Variable	Donation			
	\$0 (selfish)	\$10 (fair)	\$20 (generous)	Not given
Warmth	3.51 _a (1.31)	5.29 _b (.87)	5.17 _b (1.10)	4.83 _b (1.48)
Competence	4.90 _b (1.34)	4.89 _b (.80)	5.33 _b (1.11)	5.00 _b (1.04)
Tom's salary	\$99,643 _c	\$121,667 _c	\$96,667 _c	\$114,884 _c

Note. Means that do not share the same subscript differ at a significance of $p < .05$. Standard deviations are in parentheses.

icant, $F(12, 446) = 1.29, p = .22, \eta_p^2 = .035$. Again, generosity was not seen as a sign of incompetence.

Evaluations of warmth were of greater interest. A 4 (suggested donation) \times 5 (donation amount) between-participants analysis of variance (ANOVA) on warmth yielded main effects for suggested donation, $F(3, 446) = 9.41, p < .001, \eta_p^2 = .06$, and donation amount, $F(4, 446) = 81.86, p < .0001, \eta_p^2 = .43$, qualified by the predicted interaction, $F(12, 446) = 3.85, p < .001, \eta_p^2 = .10$. As shown in Table 2, each donation condition yields asymptotic evaluations, with relatively selfish actions (below the suggested donation) evaluated more negatively as the donation decreases, but no significant effect on evaluations as donations increase beyond the suggested amount.

Discussion

These results suggest asymptotic evaluations of prosociality. It paid to be nice and give the suggested donation, but it did not pay to be nicer and give more. Experiment 1b further demonstrates that the presence of a clearly defined prosocial standard can moderate the asymptotic pattern by determining the point at which evaluations become nonmonotonic. Going above and beyond whatever value was defined as the expected prosocial standard did not create a more positive reputation in the eyes of observers.

At first glance, these results may suggest an experimental artifact: namely, a ceiling effect on warmth evaluations. Although it would still be interesting that evaluations of modestly prosocial actions were of maximum reputational value, we think ceiling effects are an unlikely explanation of our results for two reasons. First, participants' ratings in the generous condition were not particularly close to the maximum rating of 7, nor were the average evaluations even directionally more positive as the donation

amount increased. Participants could have rated the generous person more favorably but did not. Second, participants' ratings were not significantly more positive in the fair or generous conditions than in the no information condition of Experiment 1a. There are three possible explanations: A person's reputation of warmth in the absence of any information is also showing a ceiling effect, participants inferred that Tom donated the suggested amount when it was unspecified, or our participants simply failed to infer a more positive reputation from increased generosity. Of these three, the first seems least likely. However, we address this concern in subsequent experiments by measuring different perspectives on the exchange (Experiment 2), using unbounded measures (Experiment 3), or manipulating the mechanism we believe creates insensitivity to selflessness (Experiment 4b).

Experiment 2: Anticipated and Actual Reputations

Every child learns that sharing candy is the epitome of kindness. We gave one participant (the actor) the opportunity to give candy to another participant (the target) while a third (the observer) watched. By subtle inducement, actors behaved selfishly (gave away only 1 of 10 jelly beans), fairly (gave 5 of 10), or generously (gave 9 of 10). Both observers and targets then reported their impression of the actors' warmth and competence.

In addition to a conceptual replication, this design can examine another essential aspect of reputational inference: whether people understand how their actions are evaluated by others. We suggested earlier that people may be insensitive to the magnitude of selflessness because they do not spontaneously compare varying levels against each other. Because those who choose courses of action (the actors) are likely to consider several possible outcomes and thus compare between them, we expected they would be more sensitive to their degree of selflessness and would anticipate being valued for their generosity, predicting a more monotonic pattern of evaluations than they actually receive.

Method

Participants ($N = 195$) recruited from a community sample in Chicago completed the experiment for \$2. All participated in groups of three as part of a "tasting experiment." Participants first completed a self-description questionnaire to provide some individuating information and then were randomly assigned to role.

Actors received a tray of Jelly Belly's 10 most popular flavors and were told to choose some for themselves and some to give to

Table 2
Judgments of Warmth as a Function of Suggested and Actual Donations in Experiment 1b

Suggested donation	Actual donations				
	\$0	\$5	\$10	\$20	\$50
Not stated	3.32 _a (0.78)	5.39 _b (0.95)	5.37 _b (0.76)	5.66 _b (0.58)	5.97 _b (0.87)
\$10	3.49 _a (1.18)	4.35 _c (0.77)	5.38 _b (0.78)	5.86 _b (0.65)	5.95 _b (0.64)
\$20	3.69 _a (0.85)	4.34 _c (1.02)	4.96 _c (0.88)	5.56 _b (0.87)	5.70 _b (0.63)
\$50	3.64 _a (1.44)	4.43 _c (1.11)	4.79 _c (0.98)	4.31 _c (1.05)	5.50 _b (0.74)

Note. Means that do not share the same subscript across columns differ at a significance of $p < .05$. Standard deviations are in parentheses. Cells with italicized contents are those that either meet or exceed the suggested donation amount.

the target. We experimentally manipulated whether actors gave one, five, or nine jelly beans (corresponding to increasingly selfless divisions) by suggesting that it would be a great help to us if they chose a given outcome, but that they were still completely free to choose any outcome they wished (following the standard induced compliance paradigm; Cooper, Zanna, & Taves, 1978). Actors were informed that targets would not be told how the choice was made. Actors then tasted however many jelly beans they chose from the tray, rated how much they liked each jelly bean ($-3 = \text{very bad}$; $3 = \text{very good}$), and returned the tray. Actors were then told that the target and the observer would rate them on 10 traits (the five warmth and five competence items). Actors then received targets' and observers' self-description questionnaires and predicted how each person would rate them on all 10 items.

Targets were seated in a separate room. They received the tray from the experimenter, who told them that the actor had chosen to give them [one/five/nine] of the jelly beans to taste and ate the rest. Targets were not told of the experimenter's encouragement to give a particular number, and so the actor's behavior was presented as a completely free choice. Targets then tasted and rated their jelly beans, as the actors did. Observers sat near the targets and predicted how much the target liked each jelly bean while tasting it. Targets and observers then received the actor's self-description questionnaire and evaluated his or her warmth and competence.

Results

Choice manipulation. Five actors did not follow the experimenter's suggestion, instead giving 2, 2, 4, 6, and 8 jelly beans, respectively. Because these participants violated our random assignment, we excluded them from the following analyses (results are slightly stronger if we include them in the condition closest to their choice).²

Liking. As expected, participants liked (or thought another would like) the jelly beans. Liking was significantly positive for each role, $t(59) > 5.59$, $ps < .001$, $ds > 1.02$. Actors liked their jelly beans ($M = 1.96$, $SD = .88$) more than targets did ($M = 1.22$, $SD = 1.38$), paired $t(59) = 3.65$, $p < .01$, who in turn liked their jelly beans more than observers predicted ($M = 0.83$, $SD = 1.15$), paired $t(59) = 2.64$, $p = .01$.

Reputations. We created a composite of the warmth items and a composite of the competence items.³ Evaluations of the actors' competence, as well as the actors' predicted evaluations, were not affected by the amount of jelly beans given, among either targets or observers ($F_s < 2.11$, $ps > .13$).

Again, reputations of warmth were of greater interest. As in Experiments 1a and 1b, observers rated selfish actors as less warm ($M = 3.21$, $SD = 1.11$) than fair actors ($M = 5.30$, $SD = 0.74$), $t(38) = 7.00$, $p < .001$, $d = 2.22$, but they did not rate generous actors as warmer ($M = 4.16$, $SD = 1.44$) than merely fair actors (see Figure 2). In fact, they rated generous actors as significantly less warm than fair actors, $t(38) = 3.15$, $p = .003$, $d = 0.99$, but still as significantly warmer than selfish actors, $t(38) = 2.34$, $p < .05$, $d = 0.74$. This result is consistent with the inverted-U hypothesis. However, we are reluctant to interpret this result as a general pattern because none of our other experiments replicate this pattern, nor do targets replicate this pattern in this experiment. Given these failures to replicate, we suspect this pattern resulted

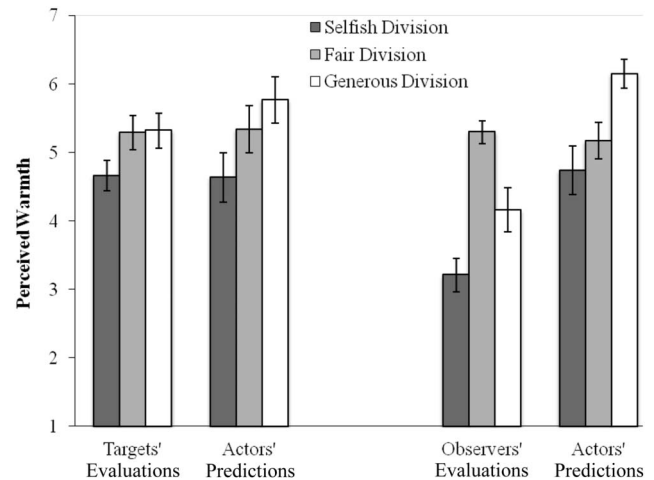


Figure 2. Predicted and actual evaluations of selfish, fair, and generous divisions in Experiment 2. Error bars represent standard errors.

from an unexpected result among the observers, who thought that targets liked their jelly beans the least. This could make our intended generous act of giving away 9 jelly beans seem significantly less generous for these observers, who might have interpreted it as shirking one's responsibility in the experiment or as simply giving away mediocre jelly beans.

Targets also rated selfish actors as marginally less warm ($M = 4.66$, $SD = 0.99$) than fair actors ($M = 5.29$, $SD = 1.13$), $t(38) = 1.87$, $p = .07$, $d = 0.59$, but they did not rate generous actors as any warmer ($M = 5.32$, $SD = 1.14$) than fair actors, $t(38) = 0.08$, $p = .93$, $d = 0.02$. These results for both observers and targets again demonstrate that selfishness comes at a reputational cost but that generosity may not produce a reputational benefit beyond equity.

Predicted reputation. Actors, in contrast, expected to be judged in a relatively linear fashion by observers for their prosocial actions, $F(2, 59) = 6.43$, $p < .01$, $\eta_p^2 = .16$. In particular, actors expected that observers would evaluate them most negatively when they behaved selfishly ($M = 4.74$, $SD = 1.61$); nonsignificantly more positively when they behaved fairly ($M = 5.17$, $SD = 1.20$), $t(38) = 0.96$, $p = .34$, $d = 0.30$; and even more positively when they behaved generously ($M = 6.15$, $SD = 0.93$), $t(38) = 2.90$, $p < .01$, $d = 0.91$. The predictions differed from observers' actual evaluations, $F(2, 57) = 8.72$, $p < .001$, $\eta_p^2 = .23$, mainly

² Among the five actors who did not follow the experimenter's suggestion, four were originally in the selfish condition (and ate 2, 4, 6, and 8 jelly beans instead of 9), and one was originally in the fair condition (and ate 2 jelly beans instead of 5). Including participants who did not follow the experimenter's subtle suggestion in the conditions that were closest to the actual amounts of jelly beans they gave did not change the reported effects in any meaningful way, except that the difference between targets' evaluations of the warmth of selfish actors and fair actors is now statistically significant, $t(41) = 2.44$, $p = .02$, $d = 0.74$.

³ The scale reliabilities were as follows. Targets' evaluations: warmth ($\alpha = .88$), competence ($\alpha = .60$); actors' predictions of targets' evaluations: warmth ($\alpha = .93$), competence ($\alpha = .69$); observers' evaluations: warmth ($\alpha = .90$), competence ($\alpha = .76$); actors' predictions of observers' evaluations: warmth ($\alpha = .93$), competence ($\alpha = .78$).

because actors did not get the reputational benefit they expected from generosity, paired $t(19) = 4.73, p < .01$.

A similar significant, albeit weaker, linear pattern emerged for the actors' predictions of the targets' impressions, $F(2, 59) = 3.16, p = .05, \eta_p^2 = .10$. Actors again expected to be judged most negatively when they behaved selfishly ($M = 4.64, SD = 1.61$); nonsignificantly more positively when they behaved fairly ($M = 5.34, SD = 1.15, t(38) = 1.59, p = .12, d = 0.50$); and nonsignificantly more positively still when they behaved generously ($M = 5.77, SD = 1.51, t(38) = 1.01, p = .32, d = 0.32$). These predictions did not, however, differ overall from the targets' actual evaluations ($F < 1, \eta_p^2 = .02$).

These results suggest three conclusions. First, they tentatively suggest that people may not completely understand how their good deeds are evaluated by others. Actors expected to be credited to some extent for their generosity by targets and observers, but were not. Second, these results cast further doubt on a ceiling effect limiting positive evaluations of selflessness. In particular, actors expected to be judged significantly more favorably by observers, and directionally by targets, when they behaved generously. If our metrics can detect differences in expected evaluations, then they are also theoretically capable of detecting differences in actual evaluations. Finally, that actors anticipated evaluations were more monotonic than tasters' actual evaluations suggests that actors were more sensitive to the scope of their prosocial actions than were the targets. This could be because actors had considered other choices they could have made but did not and were therefore more likely to evaluate their chosen option in comparison to other possible alternatives than were targets. Generous actors knew they could have been less generous but chose to follow the experimenter's suggestion instead. This comparison process for actors may have enabled them to keep their behavior in perspective more than targets and observers did. We test this possible comparison mechanism further in the Experiments 4–6.

Experiment 3: Unbounded Prosociality

Our experiments thus far used bounded evaluation measures. These measures leave open the possibility that ceiling effects contributed to our empirical results. We therefore conducted Experiment 3 with different dependent measures, including an unbounded measure of prosociality. In addition, our preceding experiments also tested only three representative actions along the continuum of actions from selfish to generous. Experiment 3 evaluated a larger number of points along this continuum.

Method

Mechanical Turk workers ($N = 182$) evaluated Bob D., a person we said came to our laboratory and received \$6 to divide between himself and another person. Participants learned that Bob gave away \$0, \$1, \$2, \$3, \$4, \$5, or \$6 (manipulated between participants). Participants then reported their impression of Bob's warmth on three different measures: how much Bob cared for the less fortunate (1 = *very little*; 7 = *very much*), the strength of Bob's selfish motivations (measured with the Selfish Motives scale, an adapted version of the MMPI Cynicism scale; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989)⁴, and how much Bob gives annually to charity (estimated in dollars, an unbounded measure of prosociality).

Results and Discussion

As shown in Table 3, evaluations were again asymptotic. On every dependent measure, Bob earned an increasingly more negative reputation as he behaved more selfishly (giving \$0, \$1, or \$2) but did not earn an increasingly more positive impression as he behaved more selflessly (giving \$4, \$5, or \$6). These effects also emerged on the unbounded measure, suggesting that ceiling effects do not explain our findings. As Bob made progressively more selfish divisions (giving \$2, \$1, or \$0), participants also predicted he gave progressively less to charity, $\beta = -.44, t(75) = 4.17, p < .001$. In contrast, when Bob made progressively more generous divisions (giving \$4, \$5, or \$6), participants did not predict that he gave significantly more to charity, $\beta = .16, t(78) = 1.39, p = .17$. Although participants could have evaluated Bob more favorably as he behaved more generously, they did not.

Experiments 4a and 4b: Single Versus Joint Evaluations

All preceding experiments suggest an asymmetry in evaluations of relatively selfish versus selfless actions. For one's reputation, it pays to be nice but pays no more to be really nice.

There are two possible explanations for this pattern. One is that the asymmetry reflects objective reputational value, such that selfishness is truly valued according to magnitude whereas selflessness is not. This could be consistent with many different interpretations of Experiments 1–3. Perhaps people value equity so highly in the contexts we have studied that a modestly prosocial action that benefits both the self and others is the ideal solution, and a distaste for inequality mitigates a more favorable impression that would come from an extremely selfless action. A selfish action, in contrast, violates both the preference for equity and prosociality, leading to more monotonic evaluations. Or, perhaps the effects we have demonstrated are consistent with Kant's (1785/2012) argument that generosity is an "imperfect duty," meaning that there is no moral imperative for a person to be generous toward others above and beyond fairness. As a result, increasingly generous actions might not be imbued with increasing reputational value. Avoiding extreme selfishness, in contrast, is a "perfect duty" because there is a clear moral imperative to avoid harming others to benefit the self. Increasing selfishness might therefore be judged with increasing moral condemnation. Whatever the particular mechanism, the reputational inferences we have observed may reflect the objective reputational value of prosocial actions.

The second explanation is that people can value increasingly selfless actions monotonically, but their intuitive judgments do not always reflect it. This may occur because people are less likely to evaluate acts of selflessness in context (by spontaneously comparing them against more or less selfless outcomes) than they are acts of selfishness (which may be more readily compared to more or less selfish outcomes). Selfishness is therefore kept in relative perspective, judged in accordance with the magnitude of selfish-

⁴ The MMPI is the Minnesota Multiphasic Personality Inventory. Due to copyright restrictions, we are not able to release the results of this measure here. However, the Selfish Motives measure reveals a pattern consistent with that of the other measures in this experiment, providing further support for our hypotheses. Further details are available by contacting either author.

Table 3
Evaluations of Varying Degrees of Prosociality in Experiment 3

Measure	Amount given						Regression slopes		
	\$0	\$1	\$2	\$3	\$4	\$5	\$6	β_{selfish}	β_{generous}
Care for others	2.04 _a	2.54 _a	3.73 _b	5.22 _c	5.91 _d	5.58 _{c,d}	5.68 _{c,d}	.51*	-.09
Gives to charity	\$126 _a	\$321 _a	\$735 _b	\$2,363 _c	\$4,059 _d	\$3,863 _{c,d}	\$2,726 _{c,d}	.44*	-.16
Log of giving	1.69 _a	4.33 _b	5.49 _b	7.16 _c	7.73 _c	7.49 _c	7.34 _c	.54*	-.10

Note. Means that do not share the same subscript within each row differ at $p < .05$. The gives to charity measure was significantly positively skewed (skewness = 2.41, $SE = .18$; kurtosis = 9.03, $SE = .36$), and so we present means from both the raw data and a log transformation of the data. The Regression slopes columns report the standardized coefficients across the relatively selfish outcomes (giving \$0, \$1, and \$2) and the relatively selfless columns (giving \$4, \$5, and \$6) for each measure, showing that evaluations are sensitive to the degree of selfishness but not to the degree of selflessness. Coefficients marked with an asterisk are significant at $p < .05$.

ness, whereas selflessness is not. As we suggested earlier, this asymmetry in comparison processes could come from the relative rarity of selfless actions compared to selfish actions. Because selfless actions are relatively rare, people lack the requisite familiarity with them to evaluate them according to their magnitude.

These two explanations differ in one essential way. The first suggests that prosociality is not valued monotonically for one of several reasons, whereas the second suggests that prosociality is valued monotonically but that judgments of selfless action may not reflect this value when evaluated in isolation. We test between these two accounts by asking participants to evaluate only one of the possible outcomes (single evaluation mode; Experiment 4a) or by evaluating multiple selfish or selfless outcomes in a fully within-participants design that makes relative comparisons explicit (joint evaluation mode; Experiment 4b). If the asymmetric pattern observed in the preceding experiments reflects objective evaluations of selfless and selfish actions, then we will observe insensitivity to the magnitude of selflessness regardless of the context of evaluation. If the asymmetric pattern reflects a difference in spontaneous comparisons, particularly a failure to judge selfless actions in context, we will observe insensitivity to the magnitude of selflessness when participants evaluate only a single selfless action but will observe sensitivity to magnitude when participants evaluate multiple outcomes.

Method

Experiment 4a. MTurk participants ($N = 427$) in the selfless condition read about an accomplished professor who received an \$80,000 grant for his research and “decided to use $\$[x]$ for his own research and to donate $\$[80,000 - x]$ to a nonprofit institution dedicated to research on poverty.”⁵ Participants in the selfish condition read about the same accomplished professor, but that he instead found a bag on the street while walking to work containing “a tall stack of \$100 bills,” totaling \$80,000. Participants then read that he “decided to take $\$[x]$ for himself and to return $\$[80,000 - x]$ to the police.” We manipulated how much the professor donated/took, from \$0 to \$80,000 in \$10,000 increments (between participants). Note that we do not compare between the selfish and generous conditions, as these scenarios differ on numerous dimensions. We instead test our hypotheses by calculating the slopes of evaluations within each of these conditions.

Because each participant in Experiment 4b had to evaluate nine outcomes, we wanted to streamline the measure of prosociality

(evaluating warmth and competence would have required each participant in Experiment 4b to answer 90 questions). Therefore, participants in both experiments simply reported how nice the action was on a scale ranging from 0 (*not nice at all*) to 100 (*very nice*).

Experiment 4b. MTurk participants ($N = 123$) completed the same procedure as 4a, except they evaluated all nine outcomes, from the most generous to the most selfish action, in random order.

Results and Discussion

Experiment 4a. As shown in Figure 3, participants evaluated the giving and taking scenarios differently. Giving nothing appeared less nice than giving \$10,000, $t(46) = 9.54$, $p < .001$, $d = 2.81$, whereas taking nothing appeared nicer than taking \$10,000, $t(46) = 6.10$, $p < .001$, $d = 1.80$.

Of greater interest were ratings of varying degrees of generosity or selfishness beyond those zero points (between \$10,000 and \$80,000). Assessing only these outcomes, participants thought that taking more money was increasingly less nice, $\beta_{\text{standardized}} = -.36$, $p < .001$, but did not think that giving more money was increasingly nicer, $\beta_{\text{standardized}} = -.08$, $p = .28$, again reflecting an asymmetric pattern ($z = 4.44$, $p < .001$). In fact, giving everything seemed no nicer than giving only one eighth of the money, $t(45) = 1.02$, $p = .31$, $d = 0.30$. Giving money had positive reputational consequences. How much the person gave was irrelevant.

Experiment 4b. Experiment 4b suggests that the asymmetric pattern comes from failing to evaluate selfless actions in relative perspective compared to selfish actions, rather than valuing all selfless actions equally (see Figure 3). In particular, when participants evaluated all possible outcomes, they continued to think that selfishly taking more money was increasingly less nice, $\beta_{\text{standardized}} = .34$, $p < .001$ (fixed-effects regression). However, they now also valued increasing selflessness, rating more generous actions as significantly nicer than less generous actions, $\beta_{\text{standardized}} = .44$, $p < .001$ (fixed-effects regression). People do value a large act of generosity over a small act, but their judgments of a single act of generosity judged in isolation do not reflect this preference because one act of generosity does not seem to call to

⁵ This scenario was inspired by a real event at our university, whereby a faculty member received a sizable financial award for his research and donated it to a university research center instead of taking it as salary or funding his own research.

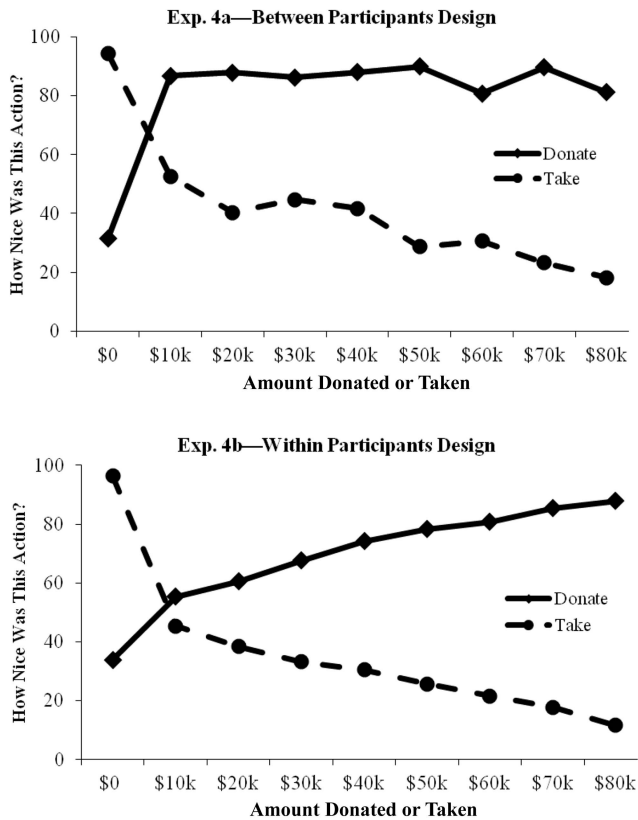


Figure 3. Evaluations of varying amounts of donated (in the research grant scenario) or taken (in the found bag scenario) in Experiments 4a (between-participants) and 4b (within-participants). Amounts are in thousands of dollars.

mind other possible acts of generosity. This suggests that the failure to value increasing degrees of prosociality comes from failing to compare increasingly prosocial actions against each other spontaneously, rather than truly valuing all of them equally.

We conducted two follow-up experiments to address potential weaknesses in Experiments 4a and 4b, both of which are described in the online supplemental materials. The first follow-up experiment manipulated the valence of the dependent measure by asking how “kind” versus “unkind” the acts of donating and taking money are. The second follow-up experiment combined Experiments 4a and 4b into a single design that enabled random assignment to either the joint or separate evaluation conditions, and it also altered the selfless scenario slightly by telling participants that the professor received an “award” rather than a “grant,” just in case participants interpreted a grant to be restricted funding. The two follow-up experiments yielded nearly identical patterns of results as observed in Experiments 4a and 4b.

Experiments 4a, 4b, and two additional follow-up experiments all demonstrate an important result: that insensitivity to increasing degrees of selflessness appears to stem from failing to judged acts of selflessness in context, rather than from truly evaluating all acts of selflessness equally. This rules out a number of alternative interpretations of Experiments 1–3 that are based the actual perceived value of selfless actions, rather than on the comparative context in which the actions are judged. These experiments suggest

that it can pay to be really nice for one’s reputation, but only if observers are reminded that you could have been a little less nice.

Experiment 5: One Versus 50 Outcomes

We suggest that reputations are sensitive to the degree of selfishness but not selflessness because the former are kept in context through comparisons whereas the latter are not. One plausible reason for this asymmetry is because selfish actions are simply more common than selfless actions (Engel, 2011) and hence are more likely to bring to mind similar comparison standards that enable more precise evaluations.

In general, the more knowledge or familiarity a person has with any stimulus, the more sensitive the person’s evaluations will be to its magnitude or precise value (Hsee & Zhang, 2010; Morewedge et al., 2009; Parducci, 1965; Stevens, 1975). Evaluations of a student’s ability, for instance, would be more sensitive to that student’s GPA than to the student’s TOEFL score because evaluators are generally more knowledgeable about GPA than about TOEFL scores. If selfish actions are more common and familiar than selfless actions, then selfless actions should be more readily evaluated in relation to their precise magnitude (like GPA) whereas selfless actions should be more insensitive to magnitude (like TOEFL scores). More important, this predicts that making people more familiar with varying degrees of a person’s selfless behavior would enable more monotonic evaluations of prosociality.

We tested this prediction in Experiment 5 by manipulating the amount of information participants had about another person. Instead of evaluating different actions by different people, as in Experiments 4a and 4b, participants evaluated a single person about whom they had little or a lot of knowledge. In particular, participants reported their impression of another person based on a single action or based on 50 actions whose mean outcome was equivalent to the single action. We predicted asymptotic evaluations of prosociality following observation of a single action, as observed in the preceding experiments, but relatively monotonic evaluations based on repeated observations because participants have more comparison information readily available to them.

Method

Participants ($N = 154$; 38% women) were recruited from Amazon.com’s Mechanical Turk and completed the experiment in exchange for a payment equivalent to an hourly rate of \$6. We used a 3 (Behavior: selfish, fair, generous) \times 2 (Knowledge: single vs. repeated observations) fully between-participants design. In the single observation condition, participants read that “Bob” came to our laboratory to complete a short experiment in which he divided \$6 between himself and a stranger (as in Experiment 3). Participants then saw a screen that reported the amount Bob kept for himself and the amount he gave to the other person. Participants saw that Bob gave \$1 (selfish), \$3 (fair), or \$5 (generous) to the other person.

In the repeated observation condition, participants read that Bob came to our lab to complete 50 short experiments in which he divided \$6 between himself and a series of strangers. Participants then saw each of Bob’s 50 ostensible choices, shown one at a time on the computer screen, in a random order within each outcome

condition (see Table 4 for complete details). The average amount given across the 50 choices was \$1, \$3, or \$5, matching the outcome observed in the single-choice condition.

When finished, participants reported their impression of Bob using the same warmth and competence scales used in Experiments 1 and 2. Finally, participants in the 50-choices condition were asked to estimate how many times they saw Bob give \$0, \$1, \$2, \$3, \$4, \$5, and \$6.

Results

Reputation. We again averaged the 5 competence ($\alpha = .69$) and 5 warmth items ($\alpha = .95$) together to create one index for evaluations of competence and one for evaluations of warmth.

Competence ratings were not significantly affected by Bob's behavior, $F(2, 153) = 1.45, p = .24$; the number of observed choices, $F(1, 153) = 2.06, p = .15$; or the interaction between them, $F(2, 153) = 0.59, p = .56$. Generosity was again not seen as a sign of incompetence.

Warmth ratings, however, were again affected by Bob's prosociality (see Figure 4). A 3 (Behavior: selfish, fair, generous) \times 2 (Observed Choices: 1 vs. 50) ANOVA revealed a main effect for behavior, $F(2, 153) = 111.68, p < .001, \eta_p^2 = .60$, and a main effect for number of actions, $F(1, 153) = 4.11, p < .05, \eta_p^2 = .03$, qualified by the predicted interaction, $F(2, 153) = 5.62, p < .01, \eta_p^2 = .07$. This interaction indicated relatively more monotonic evaluations in the repeated observation condition than in the single observation condition. In the single observation condition, participants evaluated Bob as less warm when he was selfish ($M = 3.34, SD = 1.15$) than when he was fair ($M = 5.98, SD = 0.70$), $t(52) = 9.87, p < .001, d = 2.77$, but they did not evaluate Bob differently when he was generous ($M = 5.78, SD = 0.73$) than when he was merely fair, $t(49) = 1.03, p = .31, d = 0.28$. This replicates the asymptotic pattern observed in evaluations of single actions in all of our experiments thus far.

In contrast, participants in the repeated observation condition evaluated Bob as less warm when he was selfish ($M = 3.33, SD = 1.03$) compared to fair ($M = 4.97, SD = 0.96$), $t(47) = 5.77, p < .001, d = 1.65$, and also evaluated Bob as more warm when he was generous ($M = 5.89, SD = 0.85$) compared to fair, $t(47) = 3.56, p = .001, d = 1.01$. Evaluations were sensitive to the magnitude of generosity only when participants were relative experts: when they had repeated observations about Bob's behavior to use as a basis of comparison. When participants observed Bob only once, selfishness was evaluated negatively but selflessness was not evaluated significantly more positively than merely being fair.

We made no prediction about whether the asymptotic pattern in the single-evaluation condition is due to people's overappreciation of small prosocial acts or underappreciation of large prosocial acts. In this experiment, Bob was judged as warmer when participants saw 1 fair action ($M = 5.98, SD = 0.70$) versus 50 actions that on average were fair ($M = 4.97, SD = 0.96$), $t(47) = 4.22, p < .001, d = 1.20$. Evaluations of generous and selfish behaviors did not differ between the single and repeated observation conditions ($t_s < .52, ns$), suggesting that participants placed a reputational premium on small acts of kindness judged in isolation rather than undervaluing extremely selfless actions.

Memory. In the 50-choices conditions, participants estimated the number of times they saw Bob give each of the 7 giving

amounts (\$0–\$6). Table 4 presents the relevant means broken down by giving amount. To test whether participants' memory varied by condition, we averaged the amount participants estimated Bob gave across 50 actions.⁶ In the generous condition in which Bob on average gave \$5, participants underestimated the average amount Bob actually gave ($M = \$4.60, SD = \0.56), one-sample $t(24) = 3.53, p < .01$. In the fair condition in which Bob on average gave \$3, participants' estimates did not differ from the average amount Bob actually gave ($M = \$3.03, SD = $.25$), one-sample $t(24) = 0.75, p = .46$. In the selfish condition in which Bob on average gave \$1, participants overestimated the average amount Bob actually gave ($M = \$1.93, SD = \1.53), one-sample $t(23) = 2.96, p < .01$. These patterns of recall errors suggest a slight regression to the mean in evaluations but cannot explain the relatively monotonic reputation results in the repeated observation conditions, because participants estimated the generous and selfish conditions to be less extreme than they actually were.

This experiment again demonstrates that people do evaluate generous behavior more favorably than merely fair behavior, as long as they are able to keep the action in perspective by observing multiple actions rather than a single act in isolation. In all experiments reported in this article, observers' impressions of selfish actions were relatively sensitive to the magnitude of selfishness, whereas impressions of increasingly selfless actions were insensitive to magnitude. One reason selfishness is kept in context, we suggest, is because people are more familiar with varying degrees of selfish actions in everyday life than selfless actions. Providing more information on which to base an evaluation did not alter evaluations of selfish actions, but it created more monotonic evaluations of relatively generous actions. Learning more about a person by watching variance in their behavior over time may create the knowledge base necessary to distinguish between merely nice people and extremely nice people.

In a follow-up experiment reported in the supplementary materials, we provide another test of our evaluability account by providing information about the behavioral norms of a social group rather than the history of a single individual. We find results similar to Experiment 5.

Experiment 6: Friends and Strangers

If keeping generosity in perspective requires having more observations of a person's behavior for comparison purposes, then judgments of prosociality should be more monotonic when evaluating familiar others than when evaluating unfamiliar others. Friends are more familiar than strangers, and people therefore have more observations of a friend's behavior than of a stranger's behavior. Judgments of friends are therefore more likely to be kept in a broader context than judgments of a stranger's behavior, which is likely to be judged in relative isolation. Our theory about spontaneous comparisons therefore predicts that evaluations of friends' prosociality would be more monotonic than evaluations of strangers' prosociality. Experiment 6 tested this prediction.

⁶ There can be several ways to measure the quality of recall. We also measured recall by calculating participants' mean deviations from the mean (*MAD*) for each condition, and the results were substantively similar. We believe that method we report is simpler.

Table 4
Frequencies of Divisions Shown to Participants in the 1-Action and 50-Actions Conditions in Experiment 5

Condition	Amount given							Average amount	Single-action amount
	\$0	\$1	\$2	\$3	\$4	\$5	\$6		
Generous	0	0	1	4	9	16	20	\$5.00	\$5.00
Fair	2	5	10	16	10	5	2	\$3.00	\$3.00
Selfish	20	16	9	4	1	0	0	\$1.00	\$1.00
Recall of these frequencies									
Generous	.80	1.3	2.1	6.6	8.9	13.7	16.8	\$4.60	
Fair	3.3	5.2	9.0	14.4	8.8	5.6	3.7	\$3.04	
Selfish	16.0	10.8	7.1	5.2	2.6	4.8	3.9	\$1.93	

Method

Participants ($N = 121$; 46% women) were recruited in dining halls at the University of Chicago. We used a 2 (Actor: friend vs. stranger) \times 3 (Action: selfish, fair, or generous) fully between-participants design. Experimenters approached students at dinner-time and asked if they were interested in participating in an experiment on "the psychology of interactions." Participants were recruited in pairs, with the stipulation that they had to be "friends for at least a year." Two pairs of friends participated together, creating four participants per session. We ensured that participants within pairs were friends but that participants across pairs were strangers. This enabled us to create either two pairs of friends or two pairs of strangers in each session.

The experimenter then separated the friend pairs into two separate areas in the dining hall, so that each area contained a pair of strangers. The experimenters ensured that the 2 pairs were always seated at tables that were out of sight of each other. The experimenters explained to participants that they would be randomly assigned to one of two roles, either chooser or receiver. Choosers would be given \$2 (real money) and would decide how much of this amount, if any, they would like to share with receivers. In reality, all participants were assigned to be receivers. The experimenters later on brought envelopes to participants, ostensibly from the choosers who were located in the other area in the dining

hall. The envelopes contained \$0 (selfish action), \$1 (fair), or \$2 (generous), manipulated between participants.

Experimenters told participants either that their friend was the chooser or that the stranger sitting in the other area of the dining hall was the chooser. Thus, participants thought that they received an amount of money either from their friend or from a stranger.

After seeing the amount of money left for them, participants made two assessments of the chooser. First, they rated how nice the chooser's action was on a scale ranging from 1 (*not at all*) to 7 (*very nice*). Second, they rated the chooser on dimensions of warmth of competence used in Experiments 1, 2, and 5. Of course, participants are likely to have well-formed impressions of their friends' warmth and competence, impressions that are not likely to be strongly effected by a single action in an experiment. We therefore included these measures to capture likely changes in evaluations of the stranger rather than the friend, where the impression would be based only on the one action observed in the experiment. We did not expect, consistent with basic Bayesian reasoning, that evaluations of a friend's overall character would be affected by a single action in this experiment.

At the conclusion of the experiment, participants kept the amount of money left for them in the envelopes. They were debriefed, asked not to reveal the details of this experiment to others in the dining hall, and thanked for participating.

Results

Evaluations of actions. We predicted that evaluations of a friend's action would be relatively monotonic but that evaluations of a stranger's action would be asymptotic. Table 5 presents the means for each cell in the experiment. A 2 (relationship) \times 3 (action) between-participants ANOVA on how nice the action was revealed a main effect for action, $F(2, 120) = 125.68$, $p < .001$, $\eta_p^2 = .69$; no main effect for relationship, $F(1, 120) = 0.06$, $p > .80$; and a marginally significant interaction, $F(2, 120) = 2.88$, $p = .06$, $\eta_p^2 = .05$.

Closer inspection of these results reveals the predicted pattern of evaluations. Among friends, giving \$1 was considered nicer ($M = 5.09$, $SD = 1.44$) than giving nothing ($M = 1.90$, $SD = 1.07$), $t(40) = 8.06$, $p < .001$, $d = 2.51$. More important, giving \$2 was considered nicer ($M = 6.81$, $SD = 0.68$) than giving only \$1, $t(41) = 4.95$, $p < .01$, $d = 1.53$. Evaluations of the friends' behavior was therefore relatively monotonic.

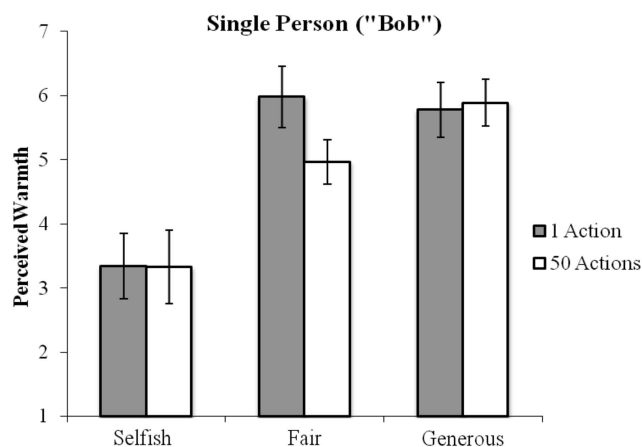


Figure 4. Evaluations of warmth as a function of prosociality and number of actions seen in Experiment 5. Error bars represent standard errors.

Table 5
Evaluations of Actions and Impressions as a Function of Relationship in Experiment 6

Type of evaluation	Relationship	Action		
		\$0 (selfish)	\$1 (fair)	\$2 (generous)
Nice action	Friend	1.94 _a	5.09 _b	6.81 _c
	Stranger	2.38 _a	5.57 _b	6.11 _b
Impression	Friend	5.11 _a	5.80 _{a,c}	6.38 _c
	Stranger	3.85 _b	5.29 _a	5.73 _{a,c}

Note. Means that do not share the same subscript within each type of evaluation differ at $p < .05$.

In contrast, strangers evaluated giving \$1 ($M = 5.57$, $SD = 1.25$) as nicer than giving nothing ($M = 2.29$, $SD = 1.69$), $t(36) = 6.88$, $p < .01$, $d = 2.21$, but they did not evaluate giving \$2 ($M = 6.10$, $SD = 1.21$) to be significantly nicer than giving only \$1, $t(39) = 1.376$, $p = .18$, $d = 0.43$. The evaluations of giving nothing or \$1 did not differ as a function of relationship ($ts < 1.20$, $ps > .25$). However, participants evaluated getting \$2 from a friend as nicer than getting \$2 from a stranger, $t(39) = 2.33$, $p = .03$, $d = 0.72$. Evaluations of strangers therefore demonstrated the same asymmetry observed in the preceding experiments, whereas evaluations of friends were more sensitive to the magnitude of prosociality. Generosity was valued more from a familiar friend than from an unknown stranger.

Reputations. We averaged the warmth items ($\alpha = .96$) and the competence items ($\alpha = .75$). A 2 (relationship) \times 3 (action) between-participants' ANOVA on competence revealed only a main effect of relationship, $F(1, 120) = 47.08$, $p < .001$, $\eta_p^2 = .29$. Not surprisingly, friends were seen as more competent ($M = 6.01$, $SD = 0.86$) than strangers ($M = 4.90$, $SD = 0.89$), $t(119) = 6.98$, $p < .01$, $d = 1.27$. More interesting, selfish strangers were evaluated as significantly more competent ($M = 5.29$, $SD = 0.70$) than fair strangers ($M = 4.74$, $SD = 0.70$), $t(36) = 2.38$, $p = .02$, but generous strangers ($M = 4.73$, $SD = 1.13$) were not seen as less competent than fair strangers, $t(39) = 0.04$, $p = .97$. Evaluations of friends' competence were not influenced by their friends' actions ($ts < .40$). Given that we do not observe this pattern of evaluations of competence among strangers in any of the other experiments discussed in this article, we do not speculate on its cause or importance further.

The warmth measures were again of greater interest. A 2 (relationship) \times 3 (action) between-participants' ANOVA on warmth revealed first a main effect of relationship, $F(1, 120) = 13.36$, $p < .01$, $\eta_p^2 = .10$, with participants rating friends ($M = 5.72$, $SD = 1.42$) more highly than strangers ($M = 4.93$, $SD = 1.44$), $t(119) = 3.02$, $p < .01$, $d = 0.55$. The ANOVA also revealed a main effect of action, $F(2, 120) = 18.49$, $p < .001$, $\eta_p^2 = .24$, and no interaction, $F(2, 120) = 0.88$, $p = .42$.

As predicted, evaluations of strangers yielded the same asymmetric effect observed in prior studies. Participants evaluated selfish strangers as less warm ($M = 3.68$, $SD = 1.45$) than fair strangers ($M = 5.29$, $SD = 0.96$), $t(36) = 4.08$, $p < .001$, $d = 1.31$, but they did not evaluate generous strangers as significantly more warm ($M = 5.63$, $SD = 1.18$) than fair strangers, $t(39) = 1.03$, $p = .31$, $d = 0.32$.

We did not expect that evaluations of a friend's general character would be significantly affected by a single action in an experiment. However, evaluations revealed a relatively small but nevertheless monotonic effect. Selfish friends were evaluated as marginally less warm ($M = 4.94$, $SD = 1.72$) than fair friends ($M = 5.80$, $SD = 1.36$), $t(40) = 1.81$, $p = .078$, $d = 0.55$, and generous friends were evaluated as marginally more warm ($M = 6.38$, $SD = 0.69$) than fair friends, $t(41) = 1.75$, $p = .087$, $d = 0.54$.

These findings complement the results of Experiments 4–5 by providing additional evidence that more knowledge about a target enables more monotonic evaluations of prosociality. Being really nice was evaluated more favorably than being merely nice for friends but not for strangers. It pays to be nice, and it can pay even more to be really nice but only when evaluated by familiar others.

General Discussion

Our experiments make the obvious point that it pays to be nice. Participants consistently evaluated those who behaved very selfishly (benefiting themselves but not others) more negatively than those who behaved fairly or equitably (benefiting themselves and others). More important, our experiments also make a nonobvious point: that it may not always pay to be really nice. Participants did not consistently evaluate a person who behaved very selflessly—benefiting others but not the self—more favorably than someone who behaved merely fairly.

This pattern emerged when single events were judged in isolation, suggesting that acts of selfishness were evaluated in context, whereas acts of selflessness were not. Experiments 4–6 make it clear that insensitivity to generosity comes not from failing to value prosociality in others, but rather from failing to compare increasingly generous actions against each other so as to keep them in a broader perspective. Evaluations were more sensitive to increasing degrees of prosociality when actions were compared against each other explicitly (Experiments 4a and 4b), when multiple actions were observed (Experiment 5), or when more was known about a target (Experiment 6). We suggest this asymmetry in evaluations of prosociality stems at least partly from the rarity of selflessness in everyday life compared to selfishness. Commonly observed actions provide a rich knowledge base of similar actions to serve as comparisons, whereas rare actions do not call to mind similar comparisons (Hsee & Zhang, 2010; Kahneman & Miller, 1986; Morewedge et al., 2009; Parducci, 1965; Stevens, 1975). If someone steals a victim's wallet on the street, the victim can easily call to mind even more selfish crimes that could have been worse. The victim might think, "It could have been worse. At least he didn't steal my car." But if someone returns a victim's wallet on the street, even more generous actions are unlikely to spring to mind as sources of comparison. The victim is unlikely to think, "It could have been better. He could have given me some spending money as well." With fewer instances of prosociality to draw on for comparison, small and large acts of kindness may be evaluated similarly. Further tests of the precise mechanisms that might underlie asymmetric evaluations of prosociality are a promising topic for future research.

Beyond these mechanisms, we believe our experiments raise four additional questions: whether these nonlinear patterns of social judgment can explain seemingly nonlinear patterns in social

behavior, whether these patterns of evaluation are universal or culture bound, why fairness seem to be overvalued in first impressions, and why we failed to identify any negative effects of excessive generosity, as other researchers have observed. We consider each, in turn.

Prosocial Thoughts Explaining Prosocial Actions?

We believe the asymmetric pattern of evaluations observed in our experiments may serve as the mechanism of several seemingly unrelated patterns of social behavior. For instance, some research suggests that people are more likely to punish selfishness in others than they are to reward selflessness. In one experiment, observers to a dictator game were given the opportunity to reward or punish dictators (either by giving them additional money or by taking away money; Almenberg et al., 2011). In the most relevant condition of this study, the dictator divided a pot of money between himself and an individual recipient. The observers punished selfish dictators who gave little money to the other person, but they did not reward selfless dictators who gave all of the money away any more than they rewarded the fair dictators who gave away only half of the money. Although this finding was incidental and not discussed in detail in that chapter, this asymmetry is precisely the one we would predict based on the social judgments we observed. If selflessness is not valued more than fairness in social judgment, then it should not be rewarded more than fairness in social behavior, either.

Another recent line of research evaluated reciprocity to selfish, fair, or generous actions (Gray et al., 2014). In these experiments, one person was on the receiving end of a selfish, fair, or generous action and was then given the opportunity to divide a pool of money (as in a dictator game) with a third person. Results consistently demonstrated that people reciprocated selfish and fair actions “in kind,” responding to a third person in a way that roughly matched how they were treated. However, participants did not reciprocate more generously after a generous division than after merely a fair division. Reciprocating equally to fair and generous actions is again what one would expect if the two are valued relatively equally in judgment.

Prosocial judgments guide prosocial behavior, and we would therefore expect the asymptotic pattern we observed here to have a wide variety of behavioral consequences in everyday life. Selfish people may be punished, but extremely generous people may not be socially rewarded any more than only modestly generous people.

How Universal Are Asymptotic Evaluations of Prosociality?

Different cultures appear to place different values on prosociality, arguably as a result of complex processes of cultural evolution and current social realities. A recent study examining public goods games in 16 countries revealed that prosociality—behaving selflessly toward others—may actually be punished rather than rewarded in some cultures (Herrmann, Thöni, & Gächter, 2008). Participants in that study had the opportunity to punish others for failing to contribute portions of their endowment for the benefit of everyone else. Whereas participants in all cultures punished extreme selfishness, participants in some cultures also punished extremely generous others.

Because of this finding, it seemed possible that the asymptotic pattern of evaluations we observed may vary across cultures. In particular, those that punish prosociality may actually evaluate it negatively. To test this possibility and to begin examining the universality of our results (Henrich, Heine, & Norenzayan, 2010), we conducted a replication of Experiment 3 with two Russian samples (Klein, Uskul, Grossman, Kraus, & Epley, 2014). In the Herrmann et al. (2008) study, the Russian sample was one of the cultures that punished prosociality the most. We recruited participants ($N = 186$) in person on a university campus in Novosibirsk and through e-mail solicitations to students at this university. After translating and back-translating the materials from Experiment 3, participants read that “Victor” came to our laboratory and received 180 rubles and was asked to divide that amount between himself and “Nikolai,” a person he had never met before. We then manipulated the amount Victor gave between participants, from completely selfish (0) to completely generous (180), and points in between at 30 ruble increments (30, 60, 90, 120, and 150). These were roughly equivalent to the dollar amounts used in Experiment 3, according to the exchange rate at the time. Participants then evaluated Nikolai on measures of warmth and competence.

A regression revealed an overall positive relationship between warmth and amount given ($\beta = .01, p < .001$). The most selfish division (giving 0 rubles) was evaluated more negatively ($M = 2.98, SD = 1.08$) than the fair division ($M = 4.91, SD = 1.06$), $t(50) = 6.47, p < .001, d = 1.80$, but the most generous division (giving 180 rubles; $M = 4.84, SD = 1.31$) was not evaluated more positively than the fair division, $t(46) = 0.20, p = .84, d = 0.06$. For generous actions (giving more than 90 rubles), perceptions of warmth did not change as the amount given increased ($\beta = -.003, p = .63$). These results replicate the pattern we observed with our American samples and suggest at least two possibilities. First, it could be that the pattern of evaluations we observed is similar across cultures, but the behavior that results from it might vary. Participants may form the same impression of a generous versus a fair person but may, for reasons unknown, nevertheless discourage extreme generosity in others. Second, it could be that we simply failed to replicate the patterns of behavior observed across cultures in the Herrmann et al. (2008) study. Further research is obviously necessary to reconcile these patterns of results. For now, we simply note that our pattern of reputational inferences seems robust, replicating in a culture that some research suggested could yield a very different pattern.

The Benefits of Fairness

Understanding how people evaluate prosociality may also shed light on the broad social processes that create a cooperative equilibrium between self-oriented and other-oriented motives. Cooperation is necessary for a flourishing social system (Bowles & Gintis, 2003; Hamilton, 1964; Trivers, 1971), but cooperating with others increases the risk of exploitation and significant personal loss. Considerable amounts of research have studied how violations of cooperation are punished (Barclay, 2004; Fehr & Schmidt, 1999; Feinberg, Willer, Stellar, & Keltner, 2012; Henrich et al., 2010), but much less has studied how social judgments could reward cooperation and kindness in a way that promotes cooperation.

The reputational premium for moderately prosocial behavior we observed in our experiments, coupled with insensitivity to generosity, could be a pattern of evaluations that promotes cooperative exchange without running the risk of exploitation that comes from truly selfless actions. Those who behave fairly quickly develop positive reputations that encourage others to approach them, without suffering the costs of extreme acts of selflessness. These reputational incentives could help explain existing norms for behaving nicely toward others, but not too nicely.

This reasoning is functional in nature and augments the views expressed in developmental and evolutionary explanations of fairness (e.g., Fehr, Bernhard, & Rockenbach, 2008). The reputational premium associated with fair actions (or minimally acceptable prosocial actions) creates an incentive for any individual to behave fairly, because fair actions optimize the ratio of reputational benefit for a given "unit" of personal cost. This may be especially true for zero-sum situations, in which the cost to givers is equivalent to the benefit to recipients, and further research can clarify whether zero-sum and non-zero-sum situations differ in this regard. Broadly, to drive cooperation, reputational incentives may favor not only punishment for selfish actions but rather strong rewards for minimally prosocial behavior. In fact, if behavior is shaped more powerfully by the rewards than by punishments (Bandura, 1971), it could be that the reward given for minimal prosociality is actually a more powerful determinant of future cooperation.

The Trouble With Generosity

The asymptotic pattern of evaluations we observed in isolated judgment is striking because of its failure to value highly generous actions. A neighbor who lends an egg at a time of baking need is no saint; nor is giving 10% of one's salary to a needy neighbor nearly as nice as giving 100%. In our experiments, participants valued selflessness only when making explicit comparisons between varying degrees of it. This implies that selflessness will be valued simply by observing more of it. We suspect, however, that valuing selflessness is more complicated than this, because there are at least three additional factors that could diminish its apparent value.

First, excessive generosity by others could lead to an unfavorable social comparison with the self and therefore to derogation as a self-defensive strategy. This is especially true when actors voice moralistic motivations for generosity, which may make observers wary of being judged (Monin et al., 2008). Although we did not find reliable evidence of an inverted-*U* pattern, as this self-defensive strategy would predict, our experiments may not have invoked the explicit self-comparisons necessary for others' generosity to seem threatening.

Second, people strongly believe that self-interest is a powerful determinant of others' behavior (Epley & Dunning, 2000; Miller & Ratner, 1998). This belief can lead to suspicion that an ulterior motive such as ingratiation is underlying generous actions or to the inference that generous actors do not really value the resources they are giving away. For example, a recent study finds that as people spend more time thinking about the motives for philanthropic behavior they come up with increasingly self-interested attributions (Critcher & Dunning, 2011). We believe this is exactly what happened among observers in Experiment 2, who evaluated generous jelly-bean givers more negatively than merely fair jelly-

bean givers. These observers did not seem to credit the givers for their generosity in this case, because they also seemed to believe that the jelly beans were not very desirable. Selfishness seems rarely discounted because it could have been a generous action in disguise. Generosity, in contrast, may be routinely discounted because of cynicism.

Third, the value of generosity may be associated with purity. This may set a particularly high bar for praise, because to be considered generous, actors may be required not only to give generously but also to avoid personal gain from their actions. A recent study finds that prosocial actors may be derogated if they incidentally accrue personal gain from their generous behavior (Lin-Healy & Small, 2013). A blemish on the motivations of generous actors can lead to diminished evaluations, whereas the evaluation of fair actors may remain unchanged even if evidence of their self-interest emerges.

These three reasons, in addition to the comparison processes we documented, may all diminish the reputational value of generosity. Nevertheless, many cultures consider selflessness to be a virtue, as evidenced in the examples of Mother Teresa and Gandhi discussed in the opening paragraph. This suggests a possible disconnect between actual evaluations of prosociality and the rhetoric commonly associated with it. One possible reason for this disconnect is that praise for generosity comes from the experience of self-sacrifice from the actor's perspective rather than from an observer's perspective. In Experiment 2, for instance, actors expected that they would be evaluated more favorably for their generosity than they actually were. Another possibility is that the praise for extreme selflessness comes from times in which it is explicitly compared against more moderate acts of kindness. Whatever the cause, those who give of themselves in the hope of being greatly appreciated may end up being sorely disappointed.

References

- Almenberg, J., Dreber, A., Apicella, C. L., & Rand, D. G. (2011). Third party reward and punishment: Group size, efficiency and public goods. In N. M. Palmetti & J. Russo (Eds.), *Psychology of punishment* (pp. 73–92). Hauppauge, NY: Nova.
- Ariely, D., & Loewenstein, G. (2000). When does duration matter in judgment and decision making? *Journal of Experimental Psychology: General*, *129*, 508–523. doi:10.1037/0096-3445.129.4.508
- Bandura, A. (1971). *Social learning theory*. New York, NY: General Learning Press.
- Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the "tragedy of the commons." *Evolution and Human Behavior*, *25*, 209–220. doi:10.1016/j.evolhumbehav.2004.04.002
- Berman, J. Z., & Small, D. A. (2012). Self-interest without selfishness: The hedonic benefit of imposed self-interest. *Psychological Science*, *23*, 1193–1199. doi:10.1177/0956797612441222
- Bowles, S., & Gintis, H. (2003). Origins of human cooperation. In P. Hammerstein (Ed.), *The genetic and cultural evolution of cooperation* (pp. 429–443). Cambridge, MA: MIT Press.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.
- Cooper, J., Zanna, M. P., & Taves, P. A. (1978). Arousal as a necessary condition for attitude change following induced compliance. *Journal of Personality and Social Psychology*, *36*, 1101–1106. doi:10.1037/0022-3514.36.10.1101

- Critcher, C. R., & Dunning, D. (2011). No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology, 47*, 1207–1213. doi:10.1016/j.jesp.2011.05.001
- Dawkins, R. (1976). *The selfish gene*. New York, NY: Oxford University Press.
- Dickens, C. (1843). *A Christmas carol*. London, England: Chapman & Hall.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14*, 583–610.
- Epley, N., & Dunning, D. (2000). Feeling “holier than thou”: Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79*, 861–875. doi:10.1037/0022-3514.79.6.861
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008, August 28). Egalitarianism in young children. *Nature, 454*, 1079–1083. doi:10.1038/nature07155
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*, 817–868. doi:10.1162/003355399556151
- Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: Reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology, 102*, 1015–1030. doi:10.1037/a0026650
- Fetchehauer, D., & Dunning, D. (2010). Why so cynical? Asymmetric feedback underlies misguided skepticism in the trustworthiness of others. *Psychological Science, 21*, 189–193. doi:10.1177/0956797609358586
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*, 77–83. doi:10.1016/j.tics.2006.11.005
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902.
- Giving USA. (2013). *Annual report of philanthropy for the year 2012*. Glenview, IL: Giving USA Foundation.
- Gray, K., Ward, A., & Norton, M. I. (2014). Paying it forward: Generalized reciprocity and the limits of generosity. *Journal of Experimental Psychology: General, 143*, 247–254. doi:10.1037/a0031047
- Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology, 7*, 1–16. doi:10.1016/0022-5193(64)90038-4
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., . . . Ziker, J. (2010, March 19). Markets, religion, community size, and the evolution of fairness and punishment. *Science, 327*, 1480–1484. doi:10.1126/science.1182238
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61–83. doi:10.1017/S0140525X0999152X
- Herrmann, B., Thöni, C., & Gächter, S. (2008, March 10). Antisocial punishment across societies. *Science, 319*, 1362–1367. doi:10.1126/science.1153808
- Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin, 125*, 576–590. doi:10.1037/0033-2909.125.5.576
- Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science, 5*, 343–355. doi:10.1177/1745691610374586
- Judd, C. M., James-Hawkins, L., & Yzerbyt, V. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology, 89*, 899–913. doi:10.1037/0022-3514.89.6.899
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*, 136–153. doi:10.1037/0033-295X.93.2.136
- Kant, I. (2012). *Groundwork of the metaphysics of morals* (rev. ed.; M. Gregor & J. Timmermann, Eds.). New York, NY: Cambridge University Press. (Original work published 1785)
- Klein, N., Uskul, A. K., Grossman, I., Kraus, A. A., & Epley, N. (2014). *Evaluations of prosociality across 7 cultures*. Unpublished manuscript, University of Chicago.
- Laming, D. R. J. (1997). *The measurement of sensation*. London, United Kingdom: Oxford University Press.
- Lin-Healy, F., & Small, D. A. (2013). Nice guys finish last and guys in last are nice: The clash between doing well and doing good. *Social Psychological and Personality Science, 4*, 692–698. doi:10.1177/1948550613476308
- Linville, P. W. (1982). The complexity–extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology, 42*, 193–211. doi:10.1037/0022-3514.42.2.193
- Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology, 74*, 53–62. doi:10.1037/0022-3514.74.1.53
- Minson, J. A., & Monin, B. (2012). Do-gooder derogation: Disparaging morally motivated minorities to defuse anticipated reproach. *Social Psychological and Personality Science, 3*, 200–207. doi:10.1177/1948550611415695
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*, 76–93. doi:10.1037/0022-3514.95.1.76
- Morewedge, C. K., Kassam, K. S., Hsee, C. K., & Caruso, E. M. (2009). Duration sensitivity depends on stimulus familiarity. *Journal of Experimental Psychology: General, 138*, 177–186. doi:10.1037/a0015219
- Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72*, 407–418. doi:10.1037/h0022602
- Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York, NY: Wiley.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35–57. doi:10.1086/406755
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science, 17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x
- Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology, 38*, 1139–1147. doi:10.1002/ejsp.549

Received May 22, 2014

Revision received August 22, 2014

Accepted September 1, 2014 ■