Moral Reputation and the Psychology of Giving:
Praise Judgments Track Personal Sacrifice Rather Than Social Good

Samuel G. B. Johnson (sgbjohnson@gmail.com)
School of Management, University of Bath, Bath, BA2 7AY UK

Abstract

Do we praise altruistic acts because they produce social benefits or because they require a personal sacrifice? On the one hand, utilitarianism demands that we maximize the social benefit of our actions, which could motivate altruistic acts. On the other hand, altruistic acts signal reputation precisely because personal sacrifice is a strong, costly signal. Consistent with the reputational account, these studies find that in the absence of reputational cues, people mainly rely on personal cost rather than social benefit when evaluating prosocial actors (Study 1). However, when reputation is known, personal cost acts as a much weaker signal and play a smaller role in moral evaluations (Study 2). We argue that these results have far-reaching implications for the psychology and philosophy of altruism, as well as practical import for charitable giving, particularly the effective altruism movement.

Keywords: Moral psychology; reputation; decision-making; prosocial behavior; altruism

Introduction

Moral philosophers, as well as our inner ethicists, often recommend altruism as an essential component of moral behavior. Altruistic acts have a dual character—an altruistic act requires a *personal cost* and produces a *social benefit*. The most plausible arguments for the morality of altruism seem to place the emphasis on the social benefit. Consequentialism tells us that we should act to produce the greatest good for the greatest number (e.g., Bentham, 1907/1789; Mill, 1998/1861), which often entails altruistic acts. For example, if you live in a rich country, you probably gain far less from $20 than would a family in a poor country, suggesting that the moral act is to donate the $20 (Singer, 2015).

Because human survival depends on coordinated social activity, we have moral intuitions which sometimes appear to track the conclusions of moral and legal philosophy (e.g., Mikhail, 2007); moreover, people sometimes behave like intuitive consequentialists, particularly when they have time to reflect (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). So perhaps our intuitive praise for altruistic agents stems from the same underlying consequentialist psychology that animates moral philosophers.

But do we really value altruistic acts in proportion to the social benefit they produce? Or do we rely primarily on the personal cost entailed? In many situations this a moot question, since personal sacrifice and social good are often highly correlated. This is true in both our ancestral environment, where presumably our moral intuitions evolved, and our modern environments where our moral intuitions guide behavior. If you spend two hours gathering berries, you gather more than if you spend one hour; if you give $100 rather than $50 to Oxfam, the charity can accomplish double. But these dimensions are not always so tightly correlated. The *effective altruism* movement focuses on maximizing the social good accomplished per dollar donated, since effectiveness varies hugely across different causes (MacAskill, 2015; Singer, 2015). For example, one can prevent blindness in dozens of children in the developing world for the cost of training one service dog in the developed world. To effective altruists, a donation's *quality* is at least as important as its *quantity*.

Do ordinary people, like effective altruists, prioritize social good over personal sacrifice in evaluating prosocial acts? We propose that, conversely, personal sacrifice usually looms larger. Many argue that our intuitive morality evolved to induce cooperative behavior (Goodwin, Piazza, & Rozin, 2014; Haidt, 2007; Nowak & Sigmund, 2005; Sperber & Baumard, 2012; Uhlmann, Pizarro, & Diermeier, 2015). Praise by one's social group rewards prosocial behaviors while blame penalizes anti-social behaviors, and these judgments reflect changes in moral reputation. For example, people blame others for harmless actions accompanied by "wicked desires" (Inbar, Pizarro, & Cushman, 2012) because such desires can signal poor moral character. Character judgments depend mainly on intentions, not outcome (Cushman, 2008), serving to track reliable individual differences in social behavior. Beliefs about moral reputation even have an identifiable neural basis (Delgado, Frank, & Phelps, 2005), speaking to their psychological fundamentality.

On this view, praise judgments flow from evidence of good moral character. Personal sacrifice is a stronger signal of character than social good for two reasons. First, personal sacrifice is under an actor's direct control, whereas social good depends partly on uncontrollable factors. One can write a check to Oxfam for any amount, but what the charity accomplishes depends on their decisions and on luck. Inferences based on personal sacrifice avoid such sources of noise. Second, personal sacrifice is directly observable, whereas social good is often unobservable. We see the number on the Oxfam check, but usually not how many people were helped.

Given that our interest here is in people's moral evaluations of prosocial behaviors, the most directly relevant literature would seem to be moral judgment. However, this research has focused primarily on factors influencing blame for negative acts rather than praise for positive acts (e.g., Cushman, 2008; Inbar et al., 2012;

Mikhail, 2007). Research on charitable giving does provide some hints, however. People view prosocial acts unfavorably when those acts also benefit the actor (Ariely, Bracha, & Meier, 2009; Barasch, Levine, Berman, & Small, 2014; Newman & Cain, 2014), and these perceptions have negative downstream consequences for actual prosocial behavior (Ariely, Bracha, & Meier, 2009). This suggests that personal sacrifice is a necessary condition for positive evaluations of prosocial behavior. However, presumably sacrifice alone is not sufficient—it seems doubtful that purely self-sacrificial acts would be seen as praiseworthy in the absence of some broader social benefit. The film *The Seventh Continent* depicts a middle-class family in modern Europe that destroys itself for no apparent reason, flushing their money down the toilet and committing suicide. To this audience these acts are puzzling and horrifying, not praiseworthy. Thus, the prior literature together with common intuition suggests that some degree of personal cost and some degree of social benefit are required for a prosocial act to be praised; indeed, this may be part of the very concept of altruism.

Both cost and benefit appear to track judgments of praise when we are comparing some versus none. But would they track praise when comparing a larger amount to a smaller amount? In prior work, highly prosocial acts are not seen as more praiseworthy than slightly prosocial acts, although, interestingly, people were sensitive to the degree of harm in assigning blame (Klein & Epley, 2014). However, the effects of prosocial benefits versus costs have not been teased apart in observer's moral evaluations—people may well be insensitive to the degree of cost as well as the degree of benefit in evaluating prosocial acts. On the actor side, people are more moved to donate by the plight of one than of many (Small, Loewenstein, & Slovic, 2007) and are largely indifferent to the number of individuals helped (Kahneman & Knetsch, 1992). Conversely, people are likelier to donate money when paired with a painful sacrifice (explaining, arguably, the prevalence of charity runs; Olivola & Shafir, 2013). These results again are suggestive of possible insensitivity to the degree of benefit, but do little to clarify how prosocial actors respond to the degree of cost. Moreover, it is unclear whether these results would generalize to moral evaluations rather than prosocial behaviors themselves or when cost versus benefit are pitted against one another directly.

Overall, prior work does not tell us whether moral judgments of altruist acts track personal sacrifice or social benefit. We know that some amount of personal sacrifice and social benefit are necessary conditions, but not whether one of these factors has an outsized influence compared to the other, when pitted against one another directly. This issue is critical to understanding the psychological basis of moral praise (utilitarian admiration vs. character signaling) and likely has implications for the design of charitable appeals.

Thus, the current studies investigate this issue by testing judgments of praise in response to charitable donations. The studies orthogonally manipulate the amount of personal sacrifice (size of donation) and social good (number of individuals helped), measuring judgments of praise and character. In Study 2, independent reputational cues are available, whereas in Study 1 they are not. When strong reputational cues attest to a donor's robust character, personal sacrifice is uninformative about moral character and therefore should not influence praise; however, sacrifice should have a large effect when other reputational cues are absent.

Study 1

Study 1 tested whether, absent further information about a person, judgments of prosocial behavior depend mainly on the degree of *personal sacrifice*, but not *social good*.

Method

A total of 598 American participants (56% female, M_{age} = 37.4) were recruited for Studies 1A and 1B through Mechanical Turk. Participants were excluded if they failed an attention check (see below; $N = 65$).

Participants read about a charitable donation benefiting people in a developing country. The charities focused on blindness, hunger, education, or disaster relief. The donations involved a *low*, *moderate*, or *high* monetary contribution (to manipulate personal sacrifice), and were *low* or *high* in effectiveness (to manipulate social good), with both manipulations between-subjects. These conditions always differed from one another by one order of magnitude (a factor of 10). For two of the vignettes, the beneficiary was an individual in the low-effectiveness condition and a small group in the high-effectiveness condition. For example:

> Julia decided to make a donation to charity. She donated [$20/$200/$2000] to a charity focused on international health. Her donation was used to cure [a child's/10 children's] blindness in Ethiopia.

For the other two vignettes, the beneficiary was a small group in the low-effectiveness condition and a large group in the high-effectiveness condition. For example:

> Rob decided to make a donation to charity. He donated [$12.50/$125/$1250] to a charity focused on disaster relief. His donation was used to provide basic shelter to [10/100] people for one month after a hurricane in Guatemala.

On the same screen, participants rated the praiseworthiness of the action ("Please rate the moral praiseworthiness of Julia's action") on a scale from 0 ("Not very praiseworthy") to 10 ("Extremely praiseworthy"), and the actor's character ("Please rate Julia's moral character") on a scale from 0 ("Ordinary moral character") to 10 ("Saint-like moral character").

After the main task, participants checked whether each of the four donation targets was mentioned in the study; participants making any incorrect answers were excluded.
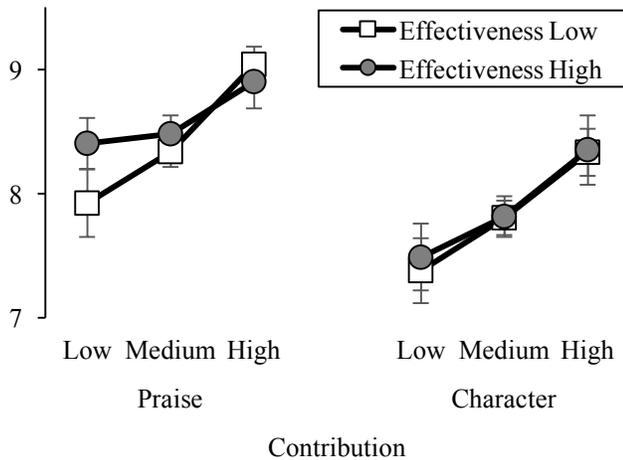
Figure 1: Results of Study 1 (character information absent). Bars represent 1 SE.

Figure 2: Results of Study 2 (character information present). Bars represent 1 SE.

Studies 1A and 1B were identical, except Study 1A included the low- and medium-contribution conditions, while Study 1B included the medium- and high-contribution conditions. Thus, both studies used 2 (personal sacrifice) x 2 (social good) designs. Given that the designs differed only in contribution levels, we combine them for analysis, to maximize statistical power and facilitate comparisons across studies.

Results

Overall, participants used the degree of personal sacrifice, but not of social good, to inform judgments of moral praise and character. The means are plotted in Figure 1.

Since contribution condition is equal-interval in log scale, it was coded as a continuous variable, (-1 = low, 0 = medium, 1 = high); effectiveness condition was contrast-coded (-1 = low, 1 = high). A linear regression was conducted, predicting moral judgments from contribution, effectiveness, and their interaction. There was a significant main effect of contribution, $b = 0.40$, $SE = 0.11$, 95% CI[0.18,0.61], $p < .001$, indicating that greater degrees of sacrifice were viewed as more morally praiseworthy. However, there was no effect of effectiveness, $b = 0.08$, $SE = 0.07$, 95% CI[-0.06,0.22], $p = .26$, nor a significant interaction, $b = -0.16$, $SE = 0.11$, 95% CI[-0.37,0.06], $p = .15$. Thus, people did not take account of social benefit in evaluating the moral praiseworthiness of the donations. Moreover, this effect did not depend on whether the less-effective donations benefited an individual or a small group: Adding this variable and its interactions to the regression model did not improve fit, $F(529,4) = 1.40$, $p = .23$. (Adding a factor for vignette also did not improve fit, indicating that there are no reliable differences in the effects across vignettes.)

The results were similar for character judgments. A regression analysis parallel to the above revealed a significant effect of contribution, $b = 0.45$, $SE = 0.12$,
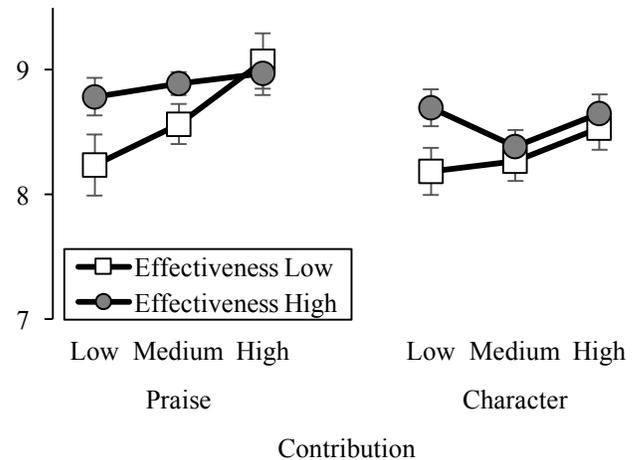
95% CI[0.21,0.69], $p < .001$, but not of effectiveness, $b = 0.02$, $SE = 0.08$, 95% CI[-0.14,0.18], $p = .82$, or the interaction between these variables, $b = -0.02$, $SE = 0.12$, 95% CI[-0.27,0.22], $p = .84$. Once again, adding the individual vs. small-group dummy-code and its interactions to the model did not improve fit, $F(4,529) = 0.62$, $p = .65$, indicating that the effect does not depend on whether the less-effective donation benefitted an individual or small group.

Discussion

Praise judgments track the amount of money sacrificed by donors, but not the social good produced by those donations. This is consistent with a signaling theory of moral praise, which assumes (i) that moral praise derives from evidence of character and (ii) that personal sacrifice is a stronger (costly, controllable, and observable) signal of character.

This mechanism will be tested more directly in Study 2. Before doing so, however, let us consider a possible boundary condition: Whether the individuals helped are closer or farther within one's moral circle (Singer, 1981). People are parochial about their charitable giving (Baron & Szymanska, 2011; see also Nagel & Waldmann, 2013), favoring causes that benefit their in-group. Perhaps altruistic acts done to benefit others in distant countries are viewed as altruistic mainly due to the signaling value (i.e., their cost), but those done to benefit one's own society are seen in a more utilitarian way. It is plausible that one would praise an altruistic act to the extent that it helped one personally, so if one identifies with one's in-group, the effectiveness of in-group help may impact praise judgments.

To test this, a replication of Study 1 was conducted (Johnson, 2018), identical except for replacing the beneficiaries living in the developed world with beneficiaries living in America (e.g., a hurricane in South Carolina rather than Guatemala). This study found a very

similar pattern of results to Study 1: Contribution was a large and robust predictor of praise, $b = 0.39$, $p < .001$, while effectiveness had a small and marginal effect, $b = 0.11$, $p = .09$. Combining the data from this follow-up with the Study 1 data, there was no significant interaction between beneficiary (in-group vs. out-group) and either contribution or effectiveness on praise. This both replicates the results of Study 1 and suggests that parochialism is not a boundary condition on the findings.

Study 2

Personal sacrifice is typically under the actor's personal control and is typically visible; therefore it can be an informative, costly signal of moral reputation. In contrast, social good is less controllable and less visible. If this drives attention to costs rather than benefits, then independent evidence of an actor's pristine moral character should decrease the relevance of individual prosocial acts for evaluating character and attenuate the effect of personal sacrifice.

Method

A total of 600 American participants (57% female, $M_{age} = 36.6$) were recruited for Study 2. Participants were excluded if they failed the same attention check used in Study 1 ($N = 91$).

Studies 2A and 2B were identical to Studies 1A and 1B, respectively, except that the vignettes were altered to include information establishing the actor's altruistic moral character. For example:

> Rob works as a receptionist, earning about $31,000 per year. He donates about 30% of his salary each year to a variety of charitable causes.

> One of the donations Rob decided to make this year was [$12.50/$125/$1250] to a charity focused on disaster relief. His donation was used to provide basic shelter to [10/100] people for one month after a hurricane in Guatemala.

The moral judgment question was rephrased so it was clear that it referred to this *specific* donation, rather than the pattern of charitable donations (e.g., "Please rate the moral praiseworthiness of Rob's [$12.50/$125/$1250] donation"). Rephrasing this question to emphasize the contribution's magnitude should, if anything, *increase* the salience of this factor, working against the hypothesis.

Results

The effects of sacrifice on perceptions of moral judgment and character were less pronounced in Study 2, when the donor's strong moral character was established, compared to Study 1, where it was not. Figure 2 plots the means.

Effects of contribution and effectiveness. Conditions were coded following the same procedure as Study 1. A linear regression was used to predict moral judgments from contribution, effectiveness, and their interaction.

For character judgments, there were no significant effects for any of the variables—neither contribution, $b = 0.07$, $SE = 0.10$, 95% CI$[-0.12, 0.26]$, $p = .48$, nor effectiveness, $b = 0.10$, $SE = 0.07$, 95% CI$[-0.03, 0.23]$, $p = .13$, nor their interaction, $b = -0.10$, $SE = 0.10$, 95% CI$[-0.29, 0.09]$, $p = .32$ reached significance. This is essentially a manipulation check, demonstrating that the manipulation successfully eliminated the diagnosticity of the specific donation for character.

For praise judgments, there was a significant effect of contribution, $b = 0.25$, $SE = 0.10$, 95% CI$[0.06, 0.44]$, $p = .009$, albeit weaker than in Study 1 (see moderated mediation analysis below). Thus, moral judgments were more positive for actors making larger contributions, but this effect was less pronounced in Study 2, where moral character was established through independent evidence, compared to Study 1.

Interestingly, there was also a modest effect of effectiveness on moral judgments, $b = 0.14$, $SE = 0.07$, 95% CI$[0.01, 0.27]$, $p = .039$, driven particularly by differences between effectiveness conditions when sacrifice was low. (This interaction, however, did not reach significance, $b = -0.16$, $SE = 0.10$, 95% CI$[-0.35, 0.03]$, $p = .10$.) This was not predicted *a priori* and should be taken with caution. One possibility is that if one is known to have a strong reputation, it may require considerable evidence to revise this default belief. When a donation is low in both magnitude and effectiveness, the combination of these two cues may provoke a negative revision to beliefs about that actor's character. A second possibility is that personal sacrifice "crowds out" social benefit when reputation is unknown, but that there is room for social benefit to play a role when there is no need to establish reputation. However, these speculations are not tested directly, and these small, unpredicted effects should be interpreted cautiously until replicated.

Moderated mediation. To test whether differences in character inferences accounted for the difference across Studies 1 and 2, a moderated mediation analysis (PROCESS Model 7; Hayes, 2013) was conducted on the combined dataset ($N = 1042$).

As shown in Figure 3, character (the mediator) was predicted by contribution, $b = 0.26$, $SE = 0.08$, $p = .001$, 95% CI$[0.11, 0.41]$ and by character information ($-1 =$ Study 1, $1 =$ Study 2), $b = 0.29$, $SE = 0.05$, $p < .001$, 95% CI$[0.19, 0.39]$. Importantly, the interaction was significant, $b = -0.19$, $SE = 0.08$, $p = .015$, 95% CI$[-0.34, -0.04]$, as contribution was a stronger predictor when character information was absent. Bootstrapping revealed that there was an indirect effect of contribution on praise judgments via character judgments for Study 1, $b = 0.23$, $SE = 0.07$, 95% CI$[0.10, 0.36]$, but not Study 2, $b = 0.03$, $SE = 0.04$, 95% CI$[-0.05, 0.12]$. This led to a significant index of moderated mediation, $b = -0.20$, $SE = 0.08$, 95% CI$[-0.35, -0.05]$. Thus, character judgments mediate the effect of contribution magnitude on praise judgments only when the actor's moral reputation is unknown.
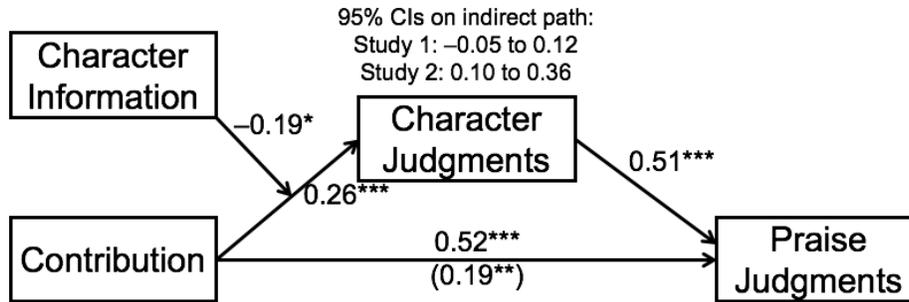
95% CIs on indirect path:
Study 1: −0.05 to 0.12
Study 2: 0.10 to 0.36

Character Information → −0.19* → Character Judgments

Contribution → 0.26*** → Character Judgments

Character Judgments → 0.51*** → Praise Judgments

Contribution → 0.52*** (0.19**) → Praise Judgments

Figure 3: Moderated mediation model for Studies 1 and 2.

Discussion

Together, Studies 1 and 2 tell a clear story about moral evaluations of prosocial acts. Without evidence of reputation, prosocial behaviors are evaluated mainly by considering their personal sacrifice, rather than the social benefit. This occurs because personal sacrifice is a controllable and visible signal of cooperativeness and thus a useful input to reputational judgments. Thus, when reputation is available, personal sacrifice is less relevant to moral evaluations

One possible concern about the character information manipulation differentiating Studies 1 and 2 is that this manipulation also introduced a reference point (the donor's salary in dollars) in addition to establishing the donor's generosity. However, since the key finding of Study 2 is that people rely *less* on personal sacrifice (in terms of dollars), it seems unlikely that this result would be explained by introducing a reference point. If anything, a reference point should make contribution amounts more salient and more readily comparable to the reference point, leading people to rely on contribution more rather than less. Nonetheless, future work might further rule out this concern by manipulating character in other ways (e.g., mentioning that the donor also volunteers her time or has dedicated her career to prosocial causes).

General Discussion

Do we admire altruists because they make personal sacrifices or because they help others? The present studies found that, for altruistic donations of money, moral praise is driven almost entirely by sacrifice (Study 1). This occurs because personal sacrifice, but not social good, is taken as a signal of moral character (Study 2).

These results have implications for the psychology, philosophy, and practice of altruism. The findings are consistent with evolutionary accounts of moral psychology, according to which our moral faculties evolved to facilitate cooperation by tracking others' reputations and creating social rewards for those willing to act for the group's benefit (Nowak & Sigmund, 2005; Sperber & Baumard, 2012). If this is true, then our evaluations (e.g., praise and blame) of prosocial behaviors would track changes to the moral reputation or character

of the actor. Since personal cost, but not social good, is usually under the actor's direct personal control, the former is a more reliable signal of cooperativeness.

How far would we expect these effects to generalize beyond this task? As discussed in conjunction with Study 1, the results do not seem to depend on the fact that the beneficiaries live in distant countries, as similar results are observed when the beneficiaries are Americans (Johnson, 2018). Other boundary conditions, however, may be plausible.

For example, the donors in the current studies may be seen as "outsourcing" the effectiveness of their charity to experts, and would thus not be seen as responsible for the outcome (e.g., Erat, 2013). This may be plausible, given previous work finding that people sometimes attribute more responsibility to individuals later in a causal chain (e.g., Brickman, Ryan, & Wortman, 1975; Spellman, 1997) as well as research documenting intransitivity beliefs about causal judgments (i.e., X causes Y and Y causes Z, but X does not cause Z; Johnson & Ahn, 2015). In that case, people may value effectiveness more when a prosocial agent contributes directly rather than indirectly. A more specific version of this possibility is that people think differently about the effectiveness of time- versus money-donations. Previous work has indeed documented differences in how people think about donations of money versus time (Johnson & Park, 2019; Liu & Aaker, 2008; Reed, Aquino, & Levy, 2007). Would effectiveness also loom larger for time-donations?

To test this, a replication of Study 1 was conducted, replacing the money-donations with time-donations (Johnson, 2018). The effects found in Study 1 were indeed reversed: Effectiveness but not sacrifice drove praise judgments. That is, unlike Study 1, contribution magnitude did not predict praise judgments, $b = 0.09$, $p = .32$, whereas effectiveness did, $b = 0.17$, $p = .007$. Thus, donation type (time vs. money) appears to be a boundary condition, such that effectiveness matters for time- but not for money-donations.

This is broadly consistent with the causal responsibility account, according to which effectiveness is only deemed irrelevant when it is outsourced to others. Indeed, low effectiveness in time-donations may signal incompetence as much as prosociality. This account alone does not

easily explain why personal sacrifice was not also used when evaluating time-donations. One possibility is that people place a greater psychological value on money than on other resources (Johnson, Zhang, & Keil, 2018), so that the sacrifice only looms large for money- but not time-donations. Future work might directly test these proposed mechanisms—the competence-signaling value of time effectiveness and the valuation difference between time and money sacrifices—in prosocial contexts.

These results are mainly bad news for effective altruism, whose *raison d'être* is improving the quality of prosocial acts, not merely their quantity. Effective altruists may receive no more social praise than *ineffective* altruists who make comparably large donations, even if the former do far more good for the world. This compounds a related problem, that people often view the importance of various causes as subjective, rather than objectively measurable (Berman et al., 2018). However, people may well be able to account for effectiveness in their moral evaluations when this factor is more salient and the causes are easily comparable. Websites like givewell.com, which directly compare charities in terms of metrics such as dollars per life saved, may be an important front on the battle for effective giving. More broadly, interventions that make both the quantity and quality of donations publicly observable may help to incentivize effective prosocial behavior.

Acknowledgements

References

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review*, *99*, 544–555.

Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology*, *107*, 393–413.

Baron, J., & Szymanska, E. (2011). Heuristics and biases in charity. In D. M. Oppenheimer & C. Y. Olivola (Eds.), *The science of giving: Experimental approaches to the study of charity* (pp. 215–235). New York, NY: Psychology Press.

Bentham, J. (1907). *An introduction to the principles of morals and legislation*. Oxford, UK: Clarendon Press. (Original work published 1789.)

Berman, J. Z., Barasch, A., Levine, E. A., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological Science*, *29*, 834–844.

Brickman, P., Ryan, K., & Wortman, C. B. (1975). Causal chains: Attribution of responsibility as a function of immediate and prior causes. *Journal of Personality and Social Psychology, 32*, 1060–1067.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*, 353–380.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, *8*, 1611–1618.

Erat, S. (2013). Avoiding lying: The case of delegated deception. *Journal of Economic Behavior & Organization*, *93*, 273–278.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*, 148–168.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144–1154.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford.

Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin*, *38*, 52–62.

Johnson, S. G. B., & Ahn, W. (2015). Causal networks or causal islands? The representation of mechanisms and the transitivity of causal judgment. *Cognitive Science, 39*, 1468–1503.

Johnson, S. G. B., & Park, S. Y. (2019). *Moral evaluations of time versus money donations.* Available at SSRN.

Johnson, S. G. B., Zhang, J., & Keil, F. C. (2018). Psychological underpinnings of zero-sum thinking. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 566–571). Austin, TX: Cognitive Science Society.

Johnson, S. G. B. (2018). *Dimensions of altruism: Do evaluations of charitable behavior track prosocial benefit or personal sacrifice?* Available at SSRN.

Kahneman, D., & Knetsch, J. (1992). Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management*, *22*, 57–70.

Klein, N., & Epley, N. (2014). The topography of generosity: Asymmetric evaluations of prosocial actions. *Journal of Experimental Psychology: General*, *143*, 2366–2379.

Liu, W., & Aaker, J. (2008). The happiness of giving: The time-ask effect. *Journal of Consumer Research*, *35*, 543–557.

MacAskill, W. (2015). *Doing good better: How effective altruism can help you make a difference*. New York,

NY: Penguin.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences*, *11*, 143–152.

Mill, J. S. (1998). *Utilitarianism*. Oxford, UK: Oxford University Press. (Original work published 1861.)

Nagel, J., & Waldmann, M. R. (2013). Deconfounding distance effects in judgments of moral obligation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 237–252.

Newman, G. E., & Cain, D. M. (2014). Tainted altruism: When doing some good is evaluated as worse than doing no good at all. *Psychological Science*, *25*, 648–655.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*, 1291–1298.

Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of Behavioral Decision Making*, *26*, 91–105.

Reed, A., Aquino, K., & Levy, E. (2007). Moral identity and judgments of charitable behaviors. *Journal of Marketing*, *71*, 178–193.

Singer, P. (1981). *The expanding circle: Ethics, evolution, and moral progress*. Princeton, NJ: Princeton University Press.

Singer, P. (2015), *The most good you can do: How effective altruism is changing ideas about living ethically*. New Haven, CT: Yale University Press.

Small, D. A., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes, 102*, 143–153.

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General, 126*, 323–348.

Sperber, D., & Baumard, N. (2012). Moral reputation: An evolutionary and cognitive perspective. *Mind & Language*, *27*, 495–518.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*, 72–81.