Journal of Happiness Studies (2022) 23:1793–1814
https://doi.org/10.1007/s10902-021-00434-w

RESEARCH PAPER

Moral Self-Appraisals Explain Emotional Rewards of Prosocial Behavior

Andrew Miles[1] · **Laura Upenieks**[2]

Accepted: 6 July 2021 / Published online: 14 July 2021

Abstract
Scholars have argued that prosocial behavior produces positive emotions because it fulfills basic psychological needs for autonomy, competence, and relatedness. These explanations have largely neglected morality, despite the fact that prosocial behaviors are widely considered to be moral. To determine whether seeing oneself as acting morally—moral self-appraisals—explained this effect, we conducted a preliminary measurement study followed by three online experiments that collectively include nearly 2000 respondents. A meta-analysis of our experimental results revealed that recalling or performing prosocial behavior has a small positive effect on positive emotion ($\beta = 0.12$, $p < 0.001$) that is partly attributable to the fact that prosocial acts encourage positive moral self-appraisals ($\beta = 0.61$, $p = 0.004$) and fulfill a psychological need for relatedness ($\beta = 0.72$, $p = 0.015$). Our results thus indicate that people feel good following prosocial behavior in part because it encourages them to view themselves as moral individuals.

Keywords Prosocial behavior · Morality · Self-determination · Mediation · Positive affect

1 Introduction

Evolutionary processes have endowed humans with innate moral capacities that begin to manifest themselves as early as infancy, and that subsequently permeate social interactions and shape how individuals see their deepest, truest selves (Haidt, 2012; Killen & Smetana, 2015; Strohminger & Nichols, 2014). These moral dispositions can prompt other-focused "prosocial" behavior which generates positive thoughts and feelings such as happiness, a sense of purpose, and gratitude (Curry et al., 2018; Dunn et al., 2014; Hofmann et al., 2014). Prosocial behavior appears to be an important avenue through which individuals can not only improve the world around them, but also enhance their emotional well-being.

Andrew Miles and Laura Upenieks authors contributed equally to this work.

✉ Andrew Miles
andrew.miles@utoronto.ca

[1] Department of Sociology, University of Toronto, 3359 Mississauga Road, Mississauga L5L 1C6, Canada

[2] Department of Sociology, Baylor University, Waco, USA

Despite the role of morality in motivating prosocial behavior, it has been largely absent from explanations of why "doing good" generates positive emotions. Instead, scholars from multiple research traditions have explained this effect as the natural result of fulfilling basic psychological needs (Dunn et al., 2014; Schwartz & Sortheix, 2018; Sheldon & Lyubomirsky, 2019). The satisfaction of basic psychological needs facilitates growth and well-being, and frustration of these needs causes psychological harm (Ryan & Deci, 2018). Scholars highlight three needs: autonomy, competence, and relatedness (hereafter "ACR needs"). Autonomy refers to a sense of volition and ownership of one's actions, while competence is a feeling of effectiveness in navigating the social environment and meeting one's goals. Relatedness refers to feeling socially connected to others. Considerable research suggests that satisfying these psychological needs produces well-being (Ryan & Deci, 2018; Weinstein & Ryan, 2010). Need satisfaction also helps explain the effects of prosocial behavior on positive emotions and happiness. Prosocial behavior has been shown to have greater impact on emotion when it enhances social connections (Aknin et al., 2013a, 2013b, 2013c; Lyubomirsky & Layous, 2013), makes a tangible difference to others (Aknin et al., 2013a, 2013b, 2013c), and when it is freely chosen (Lyubomirsky & Layous, 2013; Weinstein & Ryan, 2010). Because psychological needs are innate, a needs-based explanation indicates that the capacity to derive joy from giving might be a universal feature of human psychology (c.f., Aknin et al., 2013a, 2013b, 2013c).

Explanations based on autonomy, competence, and relatedness provide important insights into the mechanisms underlying prosocial behavior effects but overlook what is arguably the most salient feature of prosocial action: its *moral character*. Prosocial acts can be considered a subset of all moral behaviors, where moral behaviors are behaviors motivated by some type of moral conviction (Curren & Ryan, 2020). Morality is fundamentally about distinguishing between right and wrong, worthy and unworthy (Vaisey & Miles, 2014). Rightness and worth do not imply a necessary connection with ACR needs. An action might be considered morally correct even if it is not fully volitional (low autonomy), such as fulfilling a duty, or if it has little effect (low competence) and is performed anonymously (low relatedness), such as when a person donates a few dollars to an online charity. Prosocial actions can be performed in spite of inner resistance (e.g., donating one's hard-earned money to charity) and in ways that make it difficult to obtain status or other social rewards (e.g., anonymous acts of kindness). While such acts are often considered especially moral (Lambarraa & Riener, 2015), they are unlikely to promote feelings of autonomy or relatedness. Prosocial action is also frequently directed towards addressing large, complex objectives that instantiate abstract ideals such as caring for the downtrodden or protecting the weak. This can make it difficult to judge progress, which in turn could undermine feelings of competence. A person might always wonder: have I done enough?

The fact that morality is not *necessarily* connected to ACR needs should not be taken to imply that there is *never* a connection. Many moral principles are concerned with proper interactions among individuals, suggesting a close tie to relatedness. Indeed, people worldwide place moral value on interpersonal care, fairness, loyalty to social groups, and respect for legitimate hierarchical arrangements (Doğruyol et al., 2019; Haidt, 2012). Yet some types of morality do not implicate relationships in any straightforward way, indicating that morality is not fully reducible to relatedness. To cite just two examples, many people also moralize issues related to physical and metaphorical cleanliness, as well as principles like hard work that are frequently applied in non-social settings (Aquino & Reed, 2002; Haidt,

2012; Miles & Vaisey, 2015).[1] Thus, there are many ways to be moral, but only some of them involve relationships.

The fact that morality and relatedness are different concepts suggests that they may be experienced in psychologically distinct ways. Relatedness always implies a social connection, but the same is not true of morality. In fact, morality is often experienced in a deeply personal, almost self-focused way. People see morality as a fundamental part of who they truly are (Strohminger & Nichols, 2014; Strohminger et al., 2017), and respond emotionally to how well they live up to their own moral standards (Stets & Carter, 2012). Consequently, people can view themselves as moral even in circumstances where relatedness needs remain largely unfulfilled. This might occur, for instance, when a person's prosocial efforts are directed toward unknown persons (e.g., refugees in a foreign land), or are rejected.

Given these considerations, we argue that moral evaluations of the self are an important mechanism through which prosocial behavior generates positive emotion, and that these evaluations are non-redundant with ACR needs. We refer to these moral evaluations as *moral self-appraisals*. Moral self-appraisals are cognitive assessments of how closely a person is living up to his/her moral principles. Other researchers have called these evaluations "moral need satisfaction" and "moral self-image", but we argue that the term *moral self-appraisal* better describes the underlying process (Jordan et al., 2015; Prentice et al., 2019). According to Monin and Jordan (2009, p. 347), appraisals of the moral self are the answer to "one's moment to moment question of 'How moral am I?'" Thus, appraisals can vary over time and situations as people perform—or fail to perform—moral acts.[2] This makes them theoretically distinct from relatively stable constructs like moral identities and self-esteem (Jordan et al., 2015). Moral self-appraisals also differ from self-esteem in that they are cognitive evaluations rather than feelings, although these evaluations might elicit self-esteem. These theoretical distinctions have been borne out in empirical work that shows that correlations between moral self-appraisals and moral identity, self-esteem, and related constructs consistently fall below 0.30 in absolute magnitude (Jordan et al., 2015).

There are several mechanisms through which moral self-appraisals might evoke positive emotions that are non-redundant with ACR need satisfaction. First, positive moral self-appraisals might lead to positive emotions through self-enhancement processes (Dufner et al., 2019). Alternatively, because people generally hold positive views of themselves, positive moral self-appraisals might verify a person's sense of who they are, which can also generate positive emotions (Stets & Carter, 2012).[3] Prentice et al. (2019) suggest a third possibility—that humans possess a psychological need to feel moral. Moral behavior fulfills this need, and positive emotions are the natural result. It is not clear theoretically which mechanism(s) should be preferred, but for our purposes it is sufficient to note that

[1] One might argue that even seemingly non-social or weakly social moral principles have social origins, as occurs in some accounts of the evolutionary origins of morality. However, this does not imply that these moral principles *as practiced today* fulfill psychological needs for relationships. Valuing hard work, for instance, might well place strain on relationships if it translates into long hours spent in focused isolation.

[2] Moral self-appraisals are not tied to any particular type of morality, so positive feelings should occur when a person lives up to *any* moral principle, not just those associated with prosocial behavior (Prentice et al., 2020). We restrict ourselves to prosocial behavior here for consistency with past research but revisit this issue in the discussion.

[3] Self-verification theories such as affect control theory or identity theory suggest that people's emotions respond to perceptions of who one actually is.

there are multiple reasons to believe that positive moral self-appraisals will generate positive emotions.

To our knowledge, the question of whether moral self-appraisals explain the positive emotions generated by prosocial action has never been tested, though a handful of studies offer suggestive evidence. Prentice et al., (2019, 2020) have shown that moral self-appraisals predict positive emotional states even when controlling for the satisfaction of ACR needs, but do not examine whether prosocial behavior leads to moral self-appraisals. Martela and Ryan (2016a, 2016b) have demonstrated that prosocial behavior increases beneficence and that beneficence is positively associated with psychological well-being even after controlling for the satisfaction of basic psychological needs. However, these two steps in the process were examined separately, making it impossible to test for mediation. Further, beneficence is concerned with the impact one's prosocial acts, which is conceptually different from moral self-appraisals, which are judgments about the self (Martela & Ryan, 2020; Prentice et al., 2020). Collectively, these studies are consistent with the claim that moral self-appraisals help explain the effect of prosocial behavior on positive emotions, but direct evidence for the mediating role of moral self-appraisals is limited.

The primary aim of this study is to address this evidence gap. In four studies, we test whether prosocial actions generate positive emotions through moral self-appraisal processes that are non-redundant with basic psychological needs for autonomy, competence, and relatedness.

1.1 Overview of Studies

Testing the role of moral self-appraisals first requires knowing how individuals evaluate their moral character so that moral self-appraisals can be measured appropriately. Specifically, when people perform actions that they believe are moral, are they more apt to label themselves generically as "moral," or with some morally relevant characteristic like "caring" or "honest" that corresponds to the particular behavior they enacted? Study 1 addresses this question.

Studies 2 through 4 experimentally examined the role of moral self-appraisals in explaining the effect of prosocial behavior on positive emotions. All three studies shared a common format (see Fig. 2). Respondents were randomly assigned to either perform or recall a prosocial behavior, and immediately afterwards reported their current emotions. We elicited emotions without making any reference to the prosocial behavior task to encourage people to report what they were *actually* feeling rather than what they thought they *should* be feeling based on their actions. We then asked respondents to reflect on their behavior and report their moral self-appraisals and on the extent to which their behavior fulfilled needs for autonomy, competence, and relatedness.

Determining whether moral self-appraisals help explain the effect of prosocial behavior on positive emotions is fundamentally a question of mediation. Traditionally, mediation has been inferred by seeing whether the effect of a predictor variable X on an outcome Y diminishes or disappears when controlling for potential mediating variables, M (Baron & Kenny, 1986). However, Hayes (2018) points out several problems with this approach, most notably that mediation can occur even in the absence of a statistically significant main effect of X on Y. This could happen, for instance, if a single X is mediated by multiple variables such that it has both positive and negative indirect effects on Y that balance each other out. Because mediation implies that X has an indirect effect on Y through M, Hayes (2018) instead recommends demonstrating mediation by directly testing for the presence

of indirect effects. We accordingly focus on indirect effects in our analyses for studies 2, 3, and 4, though we also present all estimates from the mediation models for completeness. We also control for ACR need satisfaction to isolate mediation through moral self-appraisals from mediation through these alternate pathways.

To better account for sampling variability and the idiosyncrasies of each study, we next summarize our findings using a meta-analysis. We then compare the magnitude of the meta-analytic indirect effects of prosocial behavior through moral self-appraisals to its indirect effects through the basic psychological needs of autonomy, competence, and relatedness. This allows us to determine how important moral self-appraisals are relative to psychological needs in explaining the positive emotional consequences of prosocial behavior (Hayes, 2018).

Studies 2, 3, and 4 experimentally manipulate prosocial recall/behavior, but not moral self-appraisals, leaving open the possibility that the effect of moral self-appraisals on positive emotions is non-causal. There are two potential problems. First, it is possible that emotional reactions influence how one morally evaluates their actions, rather than the reverse as we claim. Second, it is possible that unobserved variables confound the relationship between moral self-appraisals and positive emotions. Our treatment of these issues requires familiarity with the specifics of each study, so we defer it to the discussion. To preface our conclusions, however, we argue that reverse causality and unobserved confounding are possible, but unlikely.

Prior studies of morality and need fulfillment often use items from the Basic Need Satisfaction and Frustration Scales (BNSFS; Chen et al., 2015). Our data come from a project that evolved over time and did not originally anticipate focusing on basic psychological need satisfaction to the extent that we do in this paper. Consequently, we did not become aware of the BNSFS until later in the research process, and by that time had already created short measures of need satisfaction based on the theoretical definitions of autonomy, competence, and relatedness. These measures were used for studies 1–3. We updated and expanded these measures for study 4 in consultation with the BNSFS. Supplement 5 discusses the validity of our measures in depth and demonstrates on both theoretical and empirical grounds that our measures adequately capture satisfaction of ACR needs.

All studies used paid respondents from Amazon's Mechanical Turk. Data were analyzed using multiple equation linear regressions with bootstrapped standard errors (5,000 draws) and full information maximum likelihood estimation to appropriately account for missing data (Enders, 2010). All non-dichotomous variables were standardized prior to analyses. Practically, this means that experimental effects which are operationalized using binary variables are *y*-standardized, and all other effects are fully standardized.

2 Study 1

2.1 Participants and Procedures

Study 1 was a non-experimental study aimed at determining how individuals evaluate their moral character following a prosocial act. After answering a number of questions unrelated to the present research, respondents were told that they had completed the study and would be paid $0.50. They were given the option of answering additional questions for a $1.00 reward, which 96% of respondents chose to do. These respondents were then informed that they could either keep their reward or donate it to one of four charities (chosen to appeal to

a range of interests). We regarded those who donated their reward to any charity as having acted prosocially. Eight observations missing data on all study variables were removed, leaving a final sample of N = 334. Of these, 40 donated their reward (12%).

2.2 Measures

Moral Self-Appraisals: Jordan et al. (2015) developed a scale to capture moral self-appraisals (they use the term "moral self-image"), but this scale evaluates moral traits in general and so might miss moment-by-moment variation. This scale also builds in a comparison to the person one "want[s] to be," which might not be the relevant comparison point for eliciting emotional reactions. Accordingly, we designed our own measure.[4] After completing the donation task, respondents were asked: "Based *only* on your decision to [keep/donate] the $1.00 bonus, how would you rate yourself on the following traits?" Traits were presented in random order and included helpful, caring, honest, fair, and moral, along with several distractor items. Response items were coded from 0 = "Not at all" to 7 = "Very much." Notably, this measure offers no comparison point so respondents could use whatever comparison seemed most natural to them.

Descriptive statistics and bivariate correlations for Study 1 are shown in Supplement 1, Table S1.2.

2.3 Analyses

To determine how people think of themselves following prosocial acts, we individually regressed each moral self-appraisal item on the indicator for having donated the reward using linear regression. This allowed us to determine which items best captured the types of moral evaluations individuals apply to themselves following a prosocial act. We added controls for gender, age, and household income, each of which was found to correlate with donating the reward in preliminary analyses (see Supplement 1, Sect. S1 for details).

2.4 Results and Discussion

Results are shown in Fig. 1. Respondents who donated their reward saw themselves as helpful and caring, and to a lesser degree as fair and moral.[5] Donation did not predict feeling honest.[6] This pattern of results corresponds closely the other-focused nature of prosocial acts. It is noteworthy that respondents who donated did not see themselves as strongly moral. This may be because the term "moral" is more abstract than terms like "helpful" and "caring" and more strongly suggests a positive judgment of overall character rather than an evaluation in relation to concrete prosocial acts. Whatever the reason, these results

[4] The moral need fulfillment scale created by Prentice and colleagues (2019) would have been a viable alternative, but we did not become aware of this scale until after data for all studies had been collected.

[5] Formal tests show that the difference between the donation coefficients in models for helpful and caring ratings is non-significant ($p = 0.223$), but that both of these coefficients are significantly larger than comparable coefficients in models for honest, fair, and moral ratings (all p 's < 0.001).

[6] Conducting the same analysis using data from studies 2–4 returns comparable results. After performing or recalling a prosocial act, respondents view themselves primarily as helpful (all 3 studies) and caring (2 studies), occasionally as moral (1 study), and never as fair or honest. Code to generate these results is available as part of this paper's replication materials.

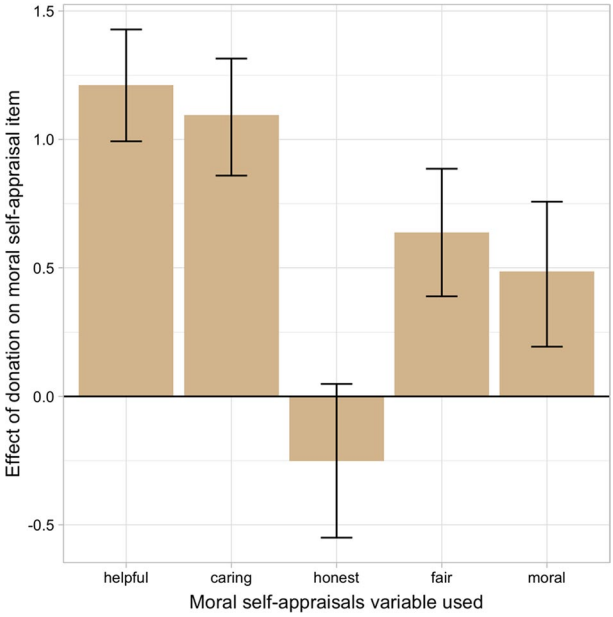

Fig. 1 Estimated effects of donating a reward on moral self-appraisal items (Study 1). Error bars are 95% confidence intervals

suggest that in the context of prosocial behavior moral self-appraisals are best captured by ratings of how helpful and caring respondents feel they are following their actions.

3 Study 2

3.1 Participants and Procedures

Study 2 operationalized prosociality using an experimental recall task widely used in work on prosocial spending and happiness. While recalling behavior is not the same as performing it, recall tasks allow respondents to relive moments and have been shown to produce similar effects on emotions (Ko et al., 2019; Strack et al., 1985).[7] A randomly selected group of respondents was asked to recall the last time they spent approximately $20 on someone else and to describe the experience in vivid detail (N = 101). Another group of respondents was randomly assigned to write about the last time they spent $20 on themselves (N = 95). We dropped seven respondents who failed attention checks and 30 who did not follow study directions, for a final sample size of N = 196.

[7] Ko et al. (2019) show that recalling and performing prosocial acts (with or without recall) have roughly equivalent effects on well-being outcomes. Recall and reflection might be responsible for producing effects even when recall is not mandated by the experimental design—it might be that individuals naturally reflect on their actions without the need for further prompting. This, however, has not been demonstrated, so it remains to be seen whether prosocial acts can have a positive influence on well-being in the absence of opportunities for reflection.

3.2 Measures

Prosocial Behavior: We created indicator variables for those assigned to the prosocial spending condition, with the self-spending recall condition as the reference category.

Positive Affect: Respondents were asked to report how happy they felt at the start of the survey (0 = very slightly or not at all; 4 = extremely). Immediately after being given the chance to donate their reward, respondents were asked to report the extent to which they were currently feeling excited, alert, active, happy, and pleased, along with a number of negative emotions, presented in random order (c.f., Aknin et al., 2013a, 2013b, 2013c). Responses ranged from 0 = "very slightly or not at all" to 4 = "extremely." Positive affect items were averaged to form a scale ($\alpha_{post} = 0.83$).

Psychological Need Fulfillment: Competence was measured by asking respondents what they hoped to accomplish by spending their money as they did, and then asking the extent to which they felt they accomplished that goal (0 = Not at all, 4 = Very much). We assessed autonomy by asking respondents how much they wanted to spend their money in the way that they had described (0 = Not at all, 4 = Very much). We measured relatedness by asking respondents the extent to which their purchase helped them feel more socially or emotionally connected to a friend, loved one, or another person (0 = Not at all, 4 = Very much).

Moral Self-Appraisals: Consistent with results from Study 1, moral self-appraisals were measured by asking respondents to think about how they had behaved (during the recalled act) and then to rate the extent to which they saw themselves as being "helpful" (0 = "Not at all" to 7 = "Very much").

Controls: We controlled for initial levels of positive affect in all models to increase the efficiency of estimates.

Note on Measurement: ACR needs and moral self-appraisals are measured using single items in both this study and Study 3. Our data include additional autonomy, relatedness, and moral self-appraisals items that could have been used to form scales, but this is not the case with competence. Because comparing effect sizes among these mediators is of interest, we opted not to use the additional items so that any differences in effect sizes could more plausibly be attributed to differences in the constructs rather than differences in measurement precision. However, results are substantively the same in either instance.

Descriptive statistics and bivariate correlations for Study 2 are shown in Supplement 2 Table S2.1.

3.3 Results and Discussion

Those in the prosocial spending condition were not more likely than those in the self-spending condition to feel positive affect ($\beta = 0.11$, $p = 0.256$; results not shown). Study 2 thus failed to replicate the finding that recalling prosocial behavior leads to positive affect. However, non-significant unmediated effects do not necessarily imply that mediation is absent (Hayes, 2018). Accordingly, we added moral self-appraisals and ACR needs to the model, which produced the first set of results in Fig. 2 (full results shown in Supplement 2, Table S2.2). Those in the prosocial spending condition were much more likely than controls to see themselves as having acted morally ($\beta = 0.73$ $p < 0.001$). Moral self-appraisals in turn predicted higher levels of positive affect, but only at a marginally significant level ($\beta = 0.12$, $p = 0.073$). Recalling prosocial spending had no effect

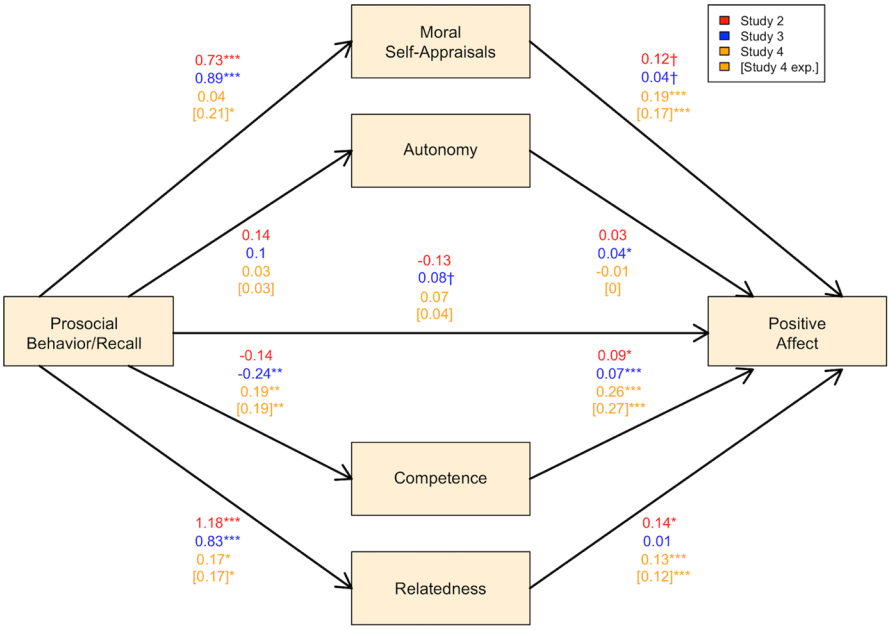

Fig. 2 Key results from studies 2, 3, 4 (pre-registered), and 4 (exploratory). Models from all studies are multiple equation linear models. Both indirect (mediation) pathways and the mediated direct effects are shown. Full-information maximum likelihood estimation used to adjust for missing data. Complete model results are presented in Supplements 2–4. p-values: ***<0.001; **<0.01; *<0.05, †<0.10 two-tailed tests. Sample sizes by study: $N_2=196$; $N_3=838$; $N_4=833$

on autonomy or competence but did predict relatedness ($\beta=1.18$ $p<0.001$). Satisfying needs for competence and relatedness predicted positive affect.

It is noteworthy that prosocial spending predicted both moral self-appraisals and relatedness, suggesting a conceptual tie to both constructs. Further, both moral self-appraisals and relatedness predicted positive affect, despite the fact that the model controlled for their shared variance. This underscores the point that moral self-appraisals—even those focused on caring behavior—are distinct from relatedness.

To test for mediation, we calculated the indirect effects of prosocial recall through moral self-appraisals and ACR needs. Following the recommendation of Hayes (2018), we used bootstrapping to generate standard errors and 95% confidence intervals. A major advantage of this approach is that it allows us to empirically approximate the sampling distribution of each indirect effect rather than assuming that it is normal, which need not be the case for indirect effects (Hayes, 2018). Non-normal sampling distributions invalidate the assumptions underlying standard hypothesis tests, which means we cannot rely on them to reliably detect "significant" mediation pathways. Instead, we

Table 1 Indirect effects from full mediation models for studies 2, 3, 4 (pre-registered), and 4 (exploratory)

	Est	S.E	95% C.I
Study 2			
Moral self-appraisals	0.09	(0.05)	[0.00, 0.20]
Competence	−0.01	(0.02)	[−0.05, 0.01]
Autonomy	0.00	(0.01)	[−0.02, 0.03]
Relatedness	0.17	(0.07)	[0.04, 0.30]
Study 3			
Moral self-appraisals	0.04	(0.02)	[0.00, 0.08]
Competence	−0.02	(0.01)	[−0.03, 0.00]
Autonomy	0.00	(0.00)	[0.00, 0.02]
Relatedness	0.01	(0.02)	[−0.03, 0.04]
Study 4			
Moral self-appraisals	0.01	(0.02)	[−0.02, 0.04]
Competence	0.05	(0.02)	[0.02, 0.09]
Autonomy	0.00	(0.00)	[0.00, 0.00]
Relatedness	0.02	(0.01)	[0.00, 0.04]
Study 4 (exploratory)			
Moral self-appraisals	0.04	(0.02)	[0.01, 0.07]
Competence	0.05	(0.02)	[0.02, 0.09]
Autonomy	0.00	(0.00)	[0.00, 0.00]
Relatedness	0.02	(0.01)	[0.00, 0.04]

Indirect effects run from prosocial behavior/recall through the variable listed to positive affect. Path coefficients for the full mediation models are shown in Fig. 2, and complete model results are presented in Supplements 2–4. Confidence intervals are based on bootstrap quantiles. Sample sizes by study: $N_2 = 196$; $N_3 = 838$; $N_4 = 833$

use confidence intervals to determine which variables mediate the prosocial recall/positive affect relationship.

Results are shown in Table 1. The largest indirect effect was through relatedness ($\beta = 0.17$), followed by moral self-appraisals ($\beta = 0.09$).[8] The confidence interval for the indirect effect through relatedness does not include 0, indicating that we can be at least 95% sure that relatedness need satisfaction mediates the relationship between prosocial recall and positive affect. The result was similar for moral self-appraisals, though here the lower bound of the confidence interval rests at 0, suggesting slightly less than 95% confidence that moral self-appraisals mediate between prosocial recall and positive affect.[9] In contrast to relatedness and moral self-appraisals, estimated indirect effects through competence and autonomy were negligible in size and had confidence intervals that straddle 0.

[8] The bootstrapped confidence interval for this difference is [-0.26, 0.10], indicating that indirect effects through moral self-appraisals and relatedness are not significantly different from one another. However, the analysis is almost certainly underpowered to detect this difference. A rough power simulation (included in the replication materials) suggests that with close to 200 respondents our power to detect a difference of 0.08 at $p < 0.05$ is less than 30%, and that achieving 80% power would require 800–900 respondents.

[9] The confidence interval for moral self-appraisals is not symmetrical around the estimated indirect effect, indicating that the sampling distribution of this effect is non-normal.

4 Study 3

4.1 Participants and Procedures

Like Study 2, Study 3 used an experimental recall task but employed a larger sample to increase the precision of estimates. Study 3 also used two comparison conditions and a different prosociality prompt that captures a wider range of prosocial behaviors. Data for Study 3 were taken from a larger study of moral behavior and emotions, but we only report results from the condition that is relevant to prosociality (plus the two comparison conditions). Respondents were asked to do the following: "Think back to a recent time that you intentionally [CONDITION SPECIFIC PROMPT]. In the space below, describe this experience as vividly and in as much detail as possible. The following questions can help you create a richer description.

What did you do? Why? Who was there? How did they respond? How did you feel?" Respondents assigned to the first condition were asked to recall a time when they had intentionally helped or cared for someone else. The second condition was our control condition, where respondents wrote about a visit to the grocery store, meant to evoke the affect associated with ordinary daily life.[10] Our third condition assigned respondents to write about a time when they purchased something for themselves that they really wanted. This condition was included to provide a comparison for the size of the prosocial effect, and to rule out the possibility that our moral self-appraisals measure captures general positivity rather than morality specifically. We dropped one duplicate response, four respondents who did not complete the experimental task (after first determining that nonresponse was not related to experimental condition), and 31 responses that were judged as non-compliant with study instructions by crowdsourced independent coders and the authors (Chmielewski & Kucker, 2019). The final sample size was N = 838, with 269 in the care condition, 273 in the control condition, and 296 in the self-indulgent purchase condition.

4.2 Measures

Prosocial Behavior: We created indicator variables for those assigned to the care and self-indulgent purchase recall conditions, with the control group as the reference category.

Positive Affect: Respondents reported their current emotions at the beginning of the survey and again immediately after the recall task. Emotion items were the same at both periods and included 10 items including excited, alert, active, happy, pleased, proud, elevated, grateful, peaceful, and content, plus a number of negative emotions. Responses were coded from 0 = "Not at all" to 8 = "Extremely." Positive emotion items were averaged to form a scale ($\alpha_{pre} = 0.93$; $\alpha_{post} = 0.94$).

Moral Self-Appraisals: Moral self-appraisals were measured as in Study 2, but the rating scale was extended to 9 scale points to capture more variation in responses (0 = "Not at all" to 8 = "Very much").

[10] People visit the grocery store for a variety of reasons, including for prosocial and self-indulgent reasons (the two comparison conditions). However, our review of the data suggests that this is not the case for most respondents. Further, either of these alternatives would elevate the average level of positive affect experienced by those in the grocery recall condition, making our analyses a conservative test of our study hypotheses.

Psychological Need Fulfillment: Measures of need satisfaction were the same as in Study 2 except for slight changes to make them fit the experimental prompts. We also extended the response scale from five scale points to nine (0 = Not at all; 8 = Very much).

Controls: The full study from which Study 3 data are taken assesses the effect of recalling different types of moral behavior on positive affect. To ensure that respondents received prompts that were relevant to their moral principles, we first sorted them into groups based on their scores on the Moral Foundations Questionnaire (MFQ30), which measures moral attitudes in five domains: care, justice, loyalty to one's ingroup, respect for authority, and personal/metaphorical purity (see Graham et al. (2011) for MFQ30 scale items). Respondents from each group were then randomly assigned to experimental conditions, but with different probabilities in each group so that respondents would only be asked to recall behaviors that corresponded to their moral principles. These probabilities were based on simulations using the anticipated distribution of MFQ30 scores, with those distributions taken from prior work using Amazon's Mechanical Turk.[11] This pre-sorting means that assignment to experimental conditions is only fully random conditional on MFQ30 scores. We accordingly controlled for MFQ30 scores in analyses to adjust for this non-random phase in the assignment process.

We also controlled for initial levels of positive affect to increase the efficiency of estimates. Descriptive statistics and bivariate correlations for Study 3 are shown in Supplement 3, Table S3.1.

4.3 Results and Discussion

Compared to controls, those in the care condition were more likely to feel positive affect after recounting their experience ($\beta = 0.12$, $p = 0.005$; results not shown). This effect was smaller than the effect of recalling a self-indulgent purchase ($\beta = 0.19$, $p < 0.001$; results not shown), though the difference in estimates was only marginally significant ($p = 0.088$). Adding moral self-appraisals and ACR needs produced the second set of results in Fig. 2 (full model given in Supplement 3, Table S3.2). Those in the care condition were much more likely than controls to see themselves as having acted morally ($\beta = 0.89$ $p < 0.001$). In contrast, those in the purchase condition reported feeling less moral, despite the fact that recalling a self-indulgent purchase was positively related to affect ($\beta = -0.29$, $p < 0.001$). This is consistent with the claim that the moral self-appraisals measure captures morality rather than general positivity. Recalling a caring act had no effect on autonomy ($\beta = 0.10$, $p = 0.263$), a strong positive effect on relatedness ($\beta = 0.83$, $p < 0.001$), and a negative effect

[11] MFQ30 subscale scores could range from 1–6. Those scoring above a 3.8 on the loyalty, authority, or purity subscales (i.e., the binding foundations) were randomly assigned to the following behavioral recall conditions, with probabilities of assignment shown in parentheses: loyalty (0.25), authority (0.25), purity (0.25), control (0.125), and self-indulgent purchase (0.125). All other respondents were assigned to conditions as follows: care (0.35), justice (0.35), control (0.15), and self-indulgent purchase (0.15). Simulations based on MFQ30 scores from a previous Amazon's Mechanical Turk sample indicated that these probabilities (combined with the indicated cutoff score) would result in a roughly equal distribution of respondents across experimental conditions.

The rationale for this assignment strategy is that most people endorse care and justice-based morality, but only some endorse morality based on ingroup loyalty, respect for authority, and purity. Consequently, anyone would likely be able to recall instances of behaving in ways that express principles care or justice, but only some respondents would be able to authentically recall behaviors that express loyalty, authority, or purity. Because MFQ30 scores were used in assigning respondents to conditions, assignment to conditions was not completely random. However, adjusting for these scores in analyses eliminates this problem.

on competence (β = −0.24, p = 0.007). This negative effect supports our argument that prosocial acts do not necessarily fulfill ACR needs and suggests that they might occasionally contribute to need frustration.

Table 1 provides direct tests of mediation, again using bootstrapped confidence intervals for inference. Recalling a caring act had the largest effect through moral self-appraisals (β = 0.04), though as in Study 2 the lower end of the confidence interval rests at 0 (95% CI: [0.00, 0.08]). The confidence interval for the indirect effect through autonomy also suggests a positive effect, though one that is likely quite small (95% CI: [0.00, 0.02]). In the case of competence this indirect effect is negative (β = −0.02, 95% CI: [−0.03, 0.00]), suggesting that the positive effect of prosocial recall occurs despite its detrimental effect on feelings of competence. Prosocial recall does not seem to be mediated by relatedness need satisfaction (β = 0.01, 95% CI: [−0.03, 0.04]).

Results from Study 3 tell the same story as Study 2: recalling one's own prosocial behavior generates positive moral self-appraisals, which in turn appear to lead to positive affect. ACR needs also played a role, but in this case the estimates were either close to 0 (autonomy) or in a negative direction (competence). Study 3 demonstrated these processes using a large experiment with high power to detect effects, and also offered evidence that moral self-appraisal effects are distinct from the effects of general positivity. Study 4 built on the strengths of Study 3 by also using an experimental design but replaced recalled behavior with real action. Study 4 also used multi-item measures of basic psychological need fulfillment and moral self-appraisals to improve the reliability of estimates.

5 Study 4

5.1 Participants and Procedures

Study 4 was a pre-registered effort to combine the merits of Studies 2 and 3 by asking respondents to perform real (not recalled) behavior under experimental conditions. We adapted the design from a prior study of prosocial behavior, basic psychological needs, and emotions (Martela & Ryan, 2016a). All respondents were asked to complete identical mathematical puzzles, but a randomly selected group was assigned to a "donation" condition in which respondents were told that the researchers would donate $0.10 to the American Red Cross for each correct answer.[12] For comparison, another randomly selected group of participants was told that they would earn a $0.10 reward for each correct answer. The total possible per person donation/reward amount was $1.20.[13] Respondents in the control condition completed the puzzles without any mention of donations or reward. To provide a stronger test of relatedness needs compared to moral self-appraisals, we allowed

[12] It might be argued that this task has only weak implications for prosociality because respondents did not choose to donate money, nor did they perform the donations themselves. However, respondents had full control over the amount of effort they invested in the task and understood that their effort would directly relate to benefits for the American Red Cross. We argue that this makes the act of trying (vs. coasting through or skipping the questions) a prosocial act. Further, even if respondents did view the task as less prosocial than otherwise might have been the case, this makes our study a conservative test of prosocial effects.

[13] The average number of correct answers (out of 12) was 6 in the control condition (SD = 3.8), 5.6 in the reward condition (SD = 3.6), and 5.8 in the donation condition (SD = 3.6). All tests of differences between the number of correct answers by condition returned p-values greater than .20.

respondents in the donation condition to dedicate their donation to a friend or loved one, if desired. After completing the puzzles, respondents were told how many items they had answered correctly.

Respondents who dropped out before being assigned to an experimental condition (N = 17) and duplicate responses (N = 7) were removed, leaving a final sample of N = 833 (N_{donate} = 282, N_{reward} = 305, $N_{control}$ = 246). The pre-registration plan can be found at https://osf.io/ea36f.

5.2 Measures

Prosocial Behavior: We created indicator variables for the donation and reward experimental conditions, with the control condition as the reference category.

Positive Affect: Respondents reported the extent to which they felt excited, happy, pleased, and content at the beginning of the study, which we averaged to form a measure of initial positive affect (α = 0.90). Responses were coded where 0 = "Not at all" and 8 = "Extremely." Immediately after completing the puzzle tasks, respondents reported their current emotions using the same 10 items as in Study 3 (α = 0.94).

Moral Self-Appraisals: Respondents were instructed to rate themselves on a list of traits based only on how they acted while completing the puzzles. We averaged self-ratings for moral, upright, principled, and good to measure moral self-appraisals (α = 0.91). Responses were coded from 0 = "Not at all" to 8 = "Very much." Note that these items capture generic morality rather than the content-specific moral terms that Study 1 suggests are more common following prosocial acts. The discrepancy arose because we had not yet considered the distinction of generic vs. content-specific moral terms at the time we pre-registered this study.

Psychological Need Fulfillment: In contrast to Studies 2 and 3, we measured psychological need fulfillment using multi-item scales. We assessed competence by first asking respondents what they hoped to accomplish by completing the numerical puzzles. We then asked them how well they believe they accomplished what they intended, how successful they were in achieving their goal, how capable they felt, and the extent to which they felt like they were completing a difficult task. All items were measured on a nine-point scale and were averaged to form a single composite measure (α = 0.88). To measure autonomy, we asked respondents to rate the extent to which they felt obligated to, wanted to, and freely chose to complete the puzzles, and the extent to which completing the puzzles expressed their true interests and values (0 = Not at all; 8 = Very much). The latter three measures were averaged to form a scale (α = 0.68). The obligation item did not load well with the others and was excluded. Relatedness was measured by asking respondents to rate the extent to which they hoped completing the puzzles would build or strengthen a relationship with a friend, loved one, or another person; the extent to which their actions actually helped them feel more socially or emotionally connected to another person; and the extent to which completing the puzzles helped them feel a sense of intimacy with another person (0 = Not at all; 8 = Very much; α = 0.95).

Controls: We controlled for the number of puzzle items respondents completely correctly and for initial levels of positive affect. We also controlled for self-reported skill at completing math-related puzzles (− 3 = Very bad; 3 = Very good) and frequency of doing math-related puzzles (0 = Never; 5 = Very often).

Descriptive statistics and bivariate correlations for Study 4 are shown in Supplement 4, Table S4.1.

5.3 Results and Discussion

Compared to controls, those in the donation condition felt more positive emotion ($\beta = 0.15$, $p = 0.012$; results not shown), as did those in the reward condition ($\beta = 0.26$, $p < 0.001$; results not shown). However, those in the reward condition did not report feeling more moral than those in the control condition ($\beta = 0.11$, $p = 0.173$; results not shown), again suggesting that moral self-appraisals are not reducible to general positivity. Adding moral self-appraisals and ACR needs produced the third set of results shown in Fig. 2 (full results shown in Supplement 4, Table S4.2). Those in the donation condition reported higher feelings of competence ($\beta = 0.19$, $p = 0.004$) and relatedness ($\beta = 0.17$, $p = 0.011$), but not autonomy ($\beta = 0.03$, $p = 0.657$). Surprisingly, those who completed puzzles for charity also did not see themselves as having acted more morally than those who simply completed puzzles without an explanation ($\beta = 0.04$, $p = 0.605$). However, those higher in moral self-appraisals still experienced more positive emotion following the puzzle task ($\beta = 0.19$, $p < 0.001$), as did those who fulfilling needs for competence ($\beta = 0.26$, $p < 0.001$) and relatedness ($\beta = 0.13$, $p < 0.001$).[14]

At first glance, the absence of a donation condition effect on moral self-appraisals is perplexing, especially considering how strong the prosocial recall/moral self-appraisal effect was in studies 2 and 3. However, our pre-registration plan specified that moral self-appraisals be measured in a content-neutral way using terms like "moral" and "principled." Study 1 indicated that people do not strongly apply these terms to themselves following prosocial behavior. Instead, they are apt to see themselves as "helpful" and "caring," terms that better reflect the concrete type of morality involved. The effects of the donation condition might have been stronger if we used an analogous measure.

To test this, we constructed a two-item measure of moral self-appraisals based on respondents' ratings of themselves as "helpful" and "caring" ($\alpha = 0.89$). We then replaced our content-neutral moral self-appraisal measure with this updated measure but kept all other analysis procedures identical to our pre-registered analyses. Key results from this exploratory analysis are shown in brackets in Fig. 2 (full results given in Supplement 4, Table S4.3). Using the updated measure, those in the donation condition reported higher moral self-appraisals ($\beta = 0.21$, $p = 0.012$). These self-appraisals in turn predicted positive affect ($\beta = 0.17$, $p < 0.001$). Table 1 shows that the indirect effect through moral self-appraisals jumps from $\beta = 0.01$ (95% CI: [-0.02, 0.04] in the pre-registered analyses to $\beta = 0.04$ (95% CI: [0.01, 0.07]) when using the improved measure. This is similar in magnitude to the indirect effect through competence ($\beta = 0.05$, 95% CI: [0.02, 0.09]) and is not significantly different from the estimate for relatedness ($\beta = 0.02$, 95% CI: [0.00, 0.04], 95% confidence interval for difference: [−0.02, 0.05]), though it exceeds the effect through autonomy ($\beta = 0.00$, 95% CI: [0.00, 0.00], 95% CI for difference: [0.01, 0.07]).

Pre-registered analyses for Study 4 were not consistent with results from studies 2 and 3, but the exploratory analyses using an alternate measure of moral self-appraisals were. This demonstrates the importance of using a self-appraisal measure appropriate to the type of moral evaluations likely to be evoked by a given behavior—in this case using terms like "caring" and "helpful" that are more directly tied to prosocial behavior than morally

[14] Our pre-registration plan also included a model showing that moral self-appraisals and psychological need fulfillment predicts positive affect, without including the donation or reward experimental conditions. This analysis does not fit with the narrative flow of our main text, but we report the results here for completeness: $\beta_{MSA} = 0.20$, $p < 0.001$; $\beta_{comp} = 0.26$, $p < 0.001$; $\beta_{auto} = -0.01$, $p = 0.735$; $\beta_{relate} = 0.121$, $p < 0.001$.

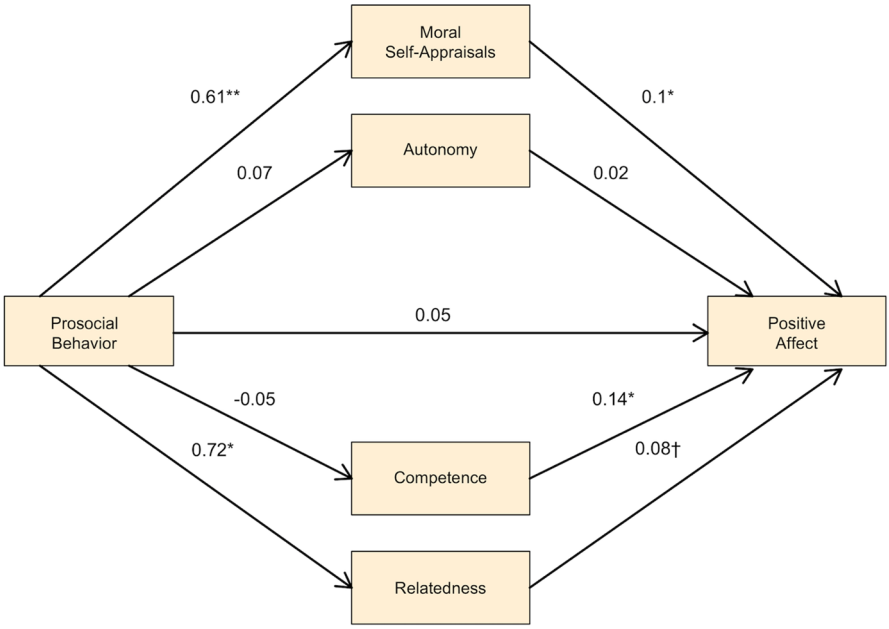

Fig. 3 Random effects meta-analyses of key results from studies 2, 3, and 4 (exploratory). *p*-values: *** < 0.001; ** < 0.01; * < 0.05, † < 0.10, two-tailed tests

generic terms like "principled" or "upright". Our results also show an unexpected positive effect of our donation condition on competence. We suspect that this is because the task we used was framed as a mathematical puzzle challenge with rewards directly tied to performance. If true, then the prosocial nature of earning money for a charity might have been accompanied—or in some cases overshadowed—by a motivation to achieve.

6 Meta-Analysis of All Mediation Analyses

Results differ somewhat across Studies 2 through 4. To give a clearer sense of the overall patterns, we conducted random-effects meta-analyses (Borenstein et al., 2009). We included the exploratory rather than pre-registered analyses from Study 4 because these used a measure of moral self-appraisals comparable to those used in Studies 2 and 3. Our meta-analyses show that those recalling or performing prosocial acts were significantly more likely to report positive emotions than controls ($\beta = 0.12$, $p < 0.001$). Adding moral self-appraisals and ACR needs gave the results shown in Fig. 3. Prosocial conditions predicted both moral self-appraisals ($\beta = 0.61$, $p = 0.004$) and relatedness ($\beta = 0.72$, $p = 0.015$), but not autonomy or competence. Of these variables, moral self-appraisals and competence predicted positive affect ($\beta = 0.10$, $p = 0.015$ and $\beta = 0.14$, $p < 0.030$, respectively), as did relatedness at a marginally significant level ($\beta = 0.08$, $p < 0.053$).

To give a clearer sense of the magnitude of the moral self-appraisal mediation pathway relative to other mechanisms, we also performed a meta-analysis of the indirect effects of prosocial recall or behavior through both moral self-appraisals and ACR needs (see

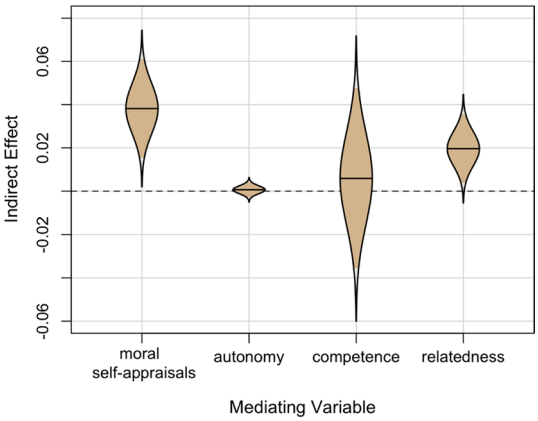

Fig. 4 Indirect effects of prosocial recall/behavior on positive affect though moral self-appraisals and ACR needs. Results are based on a meta-analysis of indirect effects from studies 2, 3, and the exploratory analyses from 4 (see Fig. 2 and Supplement tables S2.2, S3.2 and S4.2. Horizontal lines in each "cat eye" represent point estimates, and filled colored regions represent 95% confidence intervals. The width of each cat eye indicates the alternate estimates that are most plausible given the data

Table 1). Results are shown in Fig. 4. Indirect effects are displayed as "cat-eyes" with a horizontal line indicating the estimated effect, shading identifying the 95% confidence interval, and the width of the cat eye indicating the alternate estimates that are most plausible given the data (Cumming, 2014). Figure 4 reveals that prosocial behavior had the largest indirect effect through moral self-appraisals ($\beta = 0.04$, $p < 0.001$), though this estimate was not significantly different from the estimate for relatedness ($\beta = 0.02$, $p = 0.016$; $\beta_{diff} = -0.02$, $p = 0.195$). Prosocial behavior had no indirect effects through autonomy or competence.

Figure 4 underscores the point that prosocial behavior need not always fulfill ACR needs. Of the three, only relatedness needs were reliably fulfilled by prosocial recall or behavior. Further, our results demonstrate that moral self-appraisals and fulfilling relatedness needs are psychologically distinct processes, both of which help to explain the effects of prosocial acts on positive emotions. In short, the meta-analytic results suggest that people tend to feel good following prosocial behavior in part because it encourages people to view themselves as moral individuals.

7 General Discussion

We have long known that prosocial behavior can generate positive emotions (Curry et al., 2018). Our studies provide evidence that this is partially because acting prosocially or recalling prosocial acts allows individuals to see themselves as moral individuals. Further, these moral self-appraisals are not reducible to basic psychological needs for competence, autonomy, and relatedness, and are often as strong or stronger predictors of positive affect. This suggests that existing explanations of prosocial behavior's positive effects on emotions that are rooted in fulfilling basic psychological needs are accurate but incomplete (Dunn et al., 2014; Sheldon & Lyubomirsky, 2019).

An important caveat is that these conclusions are based on research designs that cannot definitively demonstrate causality for the entire mediation process. Studies 2, 3, and 4 experimentally manipulated prosocial recall/behavior, suggesting a causal effect on moral self-appraisals. However, moral self-appraisals were not experimentally manipulated, so their effects on positive emotions might be biased by unobserved confounds or endogeneity (i.e., positive emotions causing moral self-appraisals rather the reverse). We cannot rule

out these possibilities, but there are several reasons we find them unlikely. With regard to endogeneity, it seems likely people that people asked to perform or recall prosocial acts immediately recognize that engaging in such acts has moral implications. Immediate recognition means that this insight comes before prosocial acts are recalled or performed. This was particularly true of Study 4, in which respondents in the prosocial condition were necessarily aware that their efforts would help others *before* they actually completed any puzzles. Consequently, when respondents were asked to rate their morality in light of their actions, they did not need to rely on positive emotional cues to determine whether they had acted morally. This line of reasoning suggests that the causal arrow runs from moral self-appraisals to positive emotions.

What about unobserved confounds? This issue is treated at length in Supplement 6, but we summarize the main points here. First, our analytic approach makes confounding unlikely by controlling for initial positive affect in all models. To be a confound, a variable would therefore have to influence both moral self-appraisals and post-treatment affect in a way that does not run through initial affect—practically, this means that a confound would need to be generated during the course of the study. An obvious candidate is ACR need fulfillment, but we include controls for need fulfillment in all models. Of course, this only narrows the possible range of confounds. We accordingly used the approach suggested by Imai et al. (2010) to determine how sensitive our results are to the possibility of additional unobserved confounds. This technique allowed us to quantify how much confounding would be needed to eliminate the estimated indirect effects through moral self-appraisals. Our results indicated that, in theory, unobserved confounds could eliminate these effects under a range of conditions, but that these conditions are unlikely to occur in practice. To give just one example, unobserved confounds could eliminate the indirect effect in Study 3 if they explained an additional 0.8% of the variation in post-treatment affect, but they would simultaneously have to explain *all* of the remaining variation in moral self-appraisals (roughly 65%). This is possible, but unlikely. Taken together, these arguments suggest that the indirect effects we present are plausibly causal, though this issue should be addressed directly using a stronger design in future work (e.g., experimental mediation; Pirlott & MacKinnon, 2016).

Incorporating moral self-appraisals into explanations of prosocial behavior effects can be facilitated by building on insights from several areas of research. Lyubomirsky and Layous's (2013) Positive Activity Model provides a useful framework. This model predicts that positive activities can lead to emotional well-being by generating positive thoughts, a category which includes positive moral self-appraisals. The mechanisms through which moral self-appraisals or other positive thoughts produce well-being are not specified, but might include self-enhancement and/or self-verification processes (Dufner et al., 2019; Stets & Carter, 2012). The model also posits that positive activities can fulfill basic psychological needs, consistent with past work and the present analyses. However, the relationship between positive thoughts and need fulfillment is not specified. It seems plausible that feelings of competence, autonomy, or relatedness would follow conscious (or subconscious) recognition that one has achieved a goal, chosen freely, or strengthened a relationship. In this way, positive thoughts would be one mechanism that fulfills psychological needs. However, our studies found that the effects of moral self-appraisals on positive affect were not accounted for by ACR needs. Does this mean that need fulfillment cannot explain moral self-appraisal effects?

A second line of research sheds light on this issue. Prentice et al. (2019) recently proposed that individuals have a basic psychological need for morality. If this is true, then the positive effect of moral self-appraisals can be explained in the same way as the positive

effects of autonomy, competence, and relatedness—as the natural consequence of satisfying a fundamental need (Ryan & Deci, 2018). The appeal of this account is that it fits well with existing explanations of the prosocial effect on emotions. However, its validity depends on establishing that morality is a basic psychological need, a task which requires an extensive body of evidence (Ryan & Deci, 2018). Unfortunately, current evidence is sparse, and the results are mixed (Martela & Ryan, 2016b, 2020; Prentice et al., 2019, 2020). Yet pursuing this task has the potential to yield important practical benefits. If morality is a basic psychological need, then its effects on emotion should be universal. Consequently, prosocial behavior should be an effective means of promoting emotional well-being for anyone (c.f., Aknin et al., 2013a, 2013b, 2013c). Determining whether morality is a need is an important task for future research. This, in turn, will make it possible to assess whether moral self-appraisals operate through psychological need fulfillment, and how this process compares to other potential mechanisms such as self-enhancement and self-verification.

Regardless of the mechanism, our results indicate that moral self-appraisals might generate positive emotions. This has another practical implication: if seeing oneself as moral produces positive emotions, then the same emotional benefits should result from performing *any behavior* that an individual considers to be moral. Most people agree that prosocial actions are moral, but for many individuals, morality extends to concerns such as respect for proper authority; loyalty to families, communities, and other groups; and personal, sexual, and/or spiritual purity (Haidt, 2012). This offers a much wider range of behaviors that could contribute to emotional well-being, which would facilitate creating a happier, healthier population. Alternatively, the flexibility of moral emotional rewards might help explain how people become entrenched in competing moral positions that can impede the harmonious functioning of organizations, communities, and societies (Graham & Haidt, 2012). Both positive and negative consequences of these moral rewards should be examined in future work.

The results of our studies are promising, but replication is needed in different populations, and with a wider array of prosocial behaviors. A key concern will be finding a way to capture a range of behaviors tailored to the circumstances of individuals' lives—as accomplished by the recall approach used in Studies 2 and 3—while maintaining the ability to directly observe prosocial acts as in Study 4. Further research should also be directed towards distinguishing moral self-appraisals from seemingly related constructs such as moral identity, self-esteem, and interpersonal warmth. Although we did not demonstrate these conceptual in our studies, Jordan et al. (2015) offered preliminary evidence that moral self-appraisals are distinct from a variety of other constructs including moral identity and self-esteem. Similarly, Goodwin et al. (2014), while not testing moral self-appraisals directly, showed that perceptions of moral character are distinct from perceptions of warmth. However, further work testing moral self-appraisal against a broader array of constructs seems warranted.

Future research should also interrogate the boundary conditions for moral self-appraisal effects. It seems plausible that moral self-appraisal effects depend on first perceiving an act as moral, and perceptions of morality can vary individually and culturally as well across situations. Another possibility is that familiarity reduces morality effects. In the same way that completing easy tasks does less to promote a feeling of competence than finishing difficult tasks, morality effects might require moving beyond the moral acts that a person usually engages in and performing behavior that s/he perceives as atypically moral. Both of these considerations suggest a need to move beyond samples from Amazon's Mechanical Turk (AMT). Although AMT samples often return valid data, they differ from nationally representative samples in ways that might matter for work on prosociality. Research

indicates that they are less cooperative and generous (Hargittai & Shaw, 2020), and that some AMT respondents might be familiar with the standard prosociality tasks used by researchers (Peer et al., 2017; c.f., Robinson et al., 2019). For these reasons, future research that examines more types of prosocial behavior, under various conditions, and uses different samples will shed important light on the robustness of moral self-appraisals as a mechanism that explains why "doing good" is associated with "feeling good."

This research indicates that moral self-appraisals may be one route to happiness that is available to many types of people. It also raises theoretical questions about whether these effects extend to other types of moral commitments and behaviors. These questions have practical implications for understanding social conflict, and uncovering viable strategies to living a happier, more fulfilling life.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10902-021-00434-w.

Acknowledgements The authors are grateful to Gordon Brett, Miloš Brećić, Soli Dubash, Jennifer Stellar, and members of the Morality Lab at the University of Toronto for helpful comments on previous versions of this manuscript. Data and code for all analyses are available at: https://osf.io/yc7eh/.

Funding This project was supported by a Connaught New Researcher Award from the University of Toronto and a grant from the Research and Scholarly Activity Fund from the University of Toronto, Mississauga.

Declarations

Conflict of interest There are no conflicts of interest.

References

Aknin, L. B., Barrington-Leigh, C. P., Dunn, E. W., Helliwell, J. F., Burns, J., Biswas-Diener, R., Norton, M. I. (2013). Prosocial spending and well-being: Cross-cultural evidence for a psychological universal. *Journal of Personality and Social Psychology*, *104*(4), 635–652. https://doi.org/10.1037/a0031578

Aknin, L. B., Dunn, E. W., Sandstrom, G. M., & Norton, M. I. (2013b). Does social connection turn good deeds into good feelings? On the value of putting the "Social" in prosocial spending. *International Journal of Happiness and Development, 1*(2), 155–171. https://doi.org/10.1504/IJHD.2013.055643

Aknin, L. B., Dunn, E. W., Whillans, A. V., Grant, A. M., & Norton, M. I. (2013c). Making a difference matters: Impact unlocks the emotional benefits of prosocial spending. *Journal of Economic Behavior & Organization, 88*, 90–95. https://doi.org/10.1016/j.jebo.2013.01.008

Aquino, K., & Reed, A. I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423–1440. https://doi.org/10.1037/0022-3514.83.6.1423

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., & Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*(2), 216–236. https://doi.org/10.1007/s11031-014-9450-1

Chmielewski, M., & Kucker, S. C. (2019). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550619875149

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Curren, R., & Ryan, R. M. (2020). Moral self-determination: The nature, existence, and formation of moral motivation. *Journal of Moral Education, 49*(3), 295–315.

Curry, O. S., Rowland, L. A., Van Lissa, C. J., Zlotowitz, S., McAlaney, J., & Whitehouse, H. (2018). Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor. *Journal of Experimental Social Psychology*, *76*(2017), 320–329. https://doi.org/10.1016/j.jesp.2018.02.014

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences, 151*, 109547. https://doi.org/10.1016/j.paid.2019.109547

Dufner, M., Gebauer, J. E., Sedikides, C., & Denissen, J. J. A. (2019). Self-enhancement and psychological adjustment: A meta-analytic review. *Personality and Social Psychology Review, 23*(1), 48–72. https://doi.org/10.1177/1088868318756467

Dunn, E. W., Aknin, L. B., & Norton, M. I. (2014). Prosocial spending and happiness using money to benefit others pays off. *Current Directions in Psychological Science, 23*(1), 41–47. https://doi.org/10.1177/0963721413512503

Enders, C. K. (2010). *Applied missing data analysis* (1st ed.). The Guilford Press.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology, 106*(1), 148–168. https://doi.org/10.1037/a0034726

Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. Shaver & M. Mikulincer (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 11–32). APA Books.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366–385.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon.

Hargittai, E., & Shaw, A. (2020). Comparing internet experiences and prosociality in amazon mechanical Turk and population-based survey samples. *Socius, 6*, 1–11. https://doi.org/10.1177/2378023119889834

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science, 345*, 1340–1343.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science, 25*(1), 51–71. https://doi.org/10.1214/10-STS321

Jordan, J., Leliveld, M. C., & Tenbrunsel, A. E. (2015). The moral self-image scale: Measuring and understanding the malleability of the moral self. *Frontiers in Psychology*, *6*(DEC). https://doi.org/10.3389/fpsyg.2015.01878

Killen, M., & Smetana, J. G. (2015). Origins and development of morality. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (Vol. 3, pp. 701–749). Wiley-Blackwell. https://doi.org/10.1177/007542429802600109

Ko, K., Margolis, S., Revord, J., & Lyubomirsky, S. (2019). Comparing the effects of performing and recalling acts of kindness. *The Journal of Positive Psychology*. https://doi.org/10.1080/17439760.2019.1663252

Lambarraa, F., & Riener, G. (2015). On the norms of charitable giving in Islam: Two field experiments in Morocco. *Journal of Economic Behavior and Organization, 118*, 69–84. https://doi.org/10.1016/j.jebo.2015.05.006

Lyubomirsky, S., & Layous, K. (2013). How do simple positive activities increase well-being? *Current Directions in Psychological Science, 22*(1), 57–62. https://doi.org/10.1177/0963721412469809

Martela, F., & Ryan, R. M. (2016a). Prosocial behavior increases well-being and vitality even without contact with the beneficiary: Causal and behavioral evidence. *Motivation and Emotion, 40*(3), 351–357. https://doi.org/10.1007/s11031-016-9552-z

Martela, F., & Ryan, R. M. (2016b). The benefits of benevolence: Basic psychological needs, beneficence, and the enhancement of well-being. *Journal of Personality, 84*(6), 750–764. https://doi.org/10.1111/jopy.12215

Martela, F., & Ryan, R. M. (2020). Distinguishing between basic psychological needs and basic wellness enhancers: The case of beneficence as a candidate psychological need. *Motivation and Emotion, 44*(1), 116–133. https://doi.org/10.1007/s11031-019-09800-x

Miles, A., & Vaisey, S. (2015). Morality and politics: Comparing alternate theories. *Social Science Research, 53*, 252–269. https://doi.org/10.1016/j.ssresearch.2015.06.002

Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. In D. Narvaez & D. Lapsley (Eds.), *Personality, identity, and character: explorations in moral psychology* (pp. 341–354).

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Pirlott, A. G., & MacKinnon, D. P. (2016). Design approaches to experimental mediation. *Journal of Experimental Social Psychology, 66*, 29–38. https://doi.org/10.1016/j.jesp.2015.09.012

Prentice, M., Jayawickreme, E., & Fleeson, W. (2020). An experience sampling study of the momentary dynamics of moral, autonomous, competent, and related need satisfactions, moral enactments, and psychological thriving. *Motivation and Emotion*. https://doi.org/10.5005/jp/books/14190_9

Prentice, M., Jayawickreme, E., Hawkins, A., Hartley, A., Furr, R. M., & Fleeson, W. (2019). Morality as a basic psychological need. *Social Psychological and Personality Science, 10*(4), 449–460. https://doi.org/10.1177/1948550618772011

Robinson, J., Rosenzweig, C., Moss, A. J., & Litman, L. (2019). Tapped out or barely tapped? Recommendations for how to harness the vast and largely unused potential of the Mechanical Turk participant pool. *PLoS ONE, 14*(12), 1–29. https://doi.org/10.1371/journal.pone.0226394

Ryan, R. M., & Deci, E. L. (2018). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.

Schwartz, S. H., & Sortheix, F. (2018). Values and subjective well-being. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being* (pp. 1–25). Noba Scholar.

Sheldon, K. M., & Lyubomirsky, S. (2019). Revisiting the sustainable happiness model and pie chart: can happiness be successfully pursued? *The Journal of Positive Psychology*. https://doi.org/10.1080/17439760.2019.1689421

Stets, J. E., & Carter, M. J. (2012). A theory of the self for the sociology of morality. *American Sociological Review, 77*(1), 120–140. https://doi.org/10.1177/0003122411433762

Strack, F., Schwarz, N., & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology*, *49*(6), 1460–1469. https://doi.org/10.1037/0022-3514.49.6.1460

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science, 12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131*(1), 159–171.

Vaisey, S., & Miles, A. (2014). Tools from moral psychology for measuring personal moral culture. *Theory and Society, 43*(3–4), 311–332. https://doi.org/10.1007/s11186-014-9221-8

Weinstein, N., & Ryan, R. M. (2010). When helping helps: autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *Journal of Personality and Social Psychology, 98*(2), 222–244. https://doi.org/10.1037/a0016984