

Belief in altruistic motives predicts prosocial actions and inferences

Ryan W. Carlson¹ & Jamil Zaki²

¹ Department of Psychology, Yale University

² Department of Psychology, Stanford University

IN PRESS at *Psychological Reports*

Author Note

Correspondence concerning this article should be addressed to Ryan Carlson,
Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT, USA, 06511.
Email: ryan.carlson@yale.edu

Abstract

Are humans ever truly altruistic? Or are all actions, however noble, ultimately motivated by self-interest? Psychologists and philosophers have long grappled with this question, but few have considered laypeople's beliefs about the nature of prosocial motives. Here we examine these beliefs and their social correlates across two experiments (N=445). We find that people tend to believe humans can be, and frequently are, altruistically motivated—echoing prior work. Moreover, people who more strongly believe in altruistic motives act more prosocially themselves—for instance, sacrificing greater amounts of money and time to help others—a relationship that holds even when controlling for trait empathy. People who believe in altruistic motives also judge other prosocial agents to be more genuinely kind, especially when agents' motives are ambiguous. Lastly, people independently show a self-serving bias—believing their own motives for prosociality are more often altruistic than others'. Overall, this work suggests that believing in altruistic motives predicts the extent to which people both see altruism and *act* prosocially, possibly reflecting the self-fulfilling nature of such lay theories.

Keywords: altruism, self-interest, lay theories, prosocial behavior, social cognition

Belief in altruistic motives predicts prosocial actions and inferences

After seeing the philosopher Thomas Hobbes give alms to a beggar, a companion asked him to explain his action. Hobbes replied, “I was in pain to consider the miserable condition of the old man; and now my alms, giving him some relief, doth also ease me” (Mansbridge, 1990, pp. 140). Hobbes’ attributed his prosocial action to a *self-oriented* motive: improving his own mood. In fact, Hobbes famously argued that neither he nor anyone else can ever act out of *altruism*—with the ultimate goal to benefit others. To him, all kindness reflected veiled self-interest (Hobbes, 1651, Ch. 15). In the intervening centuries, psychologists have held the same position (Brown; 1986a; Cialdini, 1981; Freud, 1930; Skinner, 1978), and self-interested explanations for seemingly altruistic behavior have continued to emerge within psychology and behavioral economics as well (Andreoni & Rao, 2011; Battigalli and Dufwenberg, 2007; Bodner and Prelec, 2003; Dana et al., 2007; Maner et al., 2002; see also Carlson, Adkins, et al., 2020).

Other philosophers argue that altruism does exist (Hume, 1902; Rousseau, 1754; Smith, 1759), and both psychologists and behavioral economists have also provided compelling evidence to back this view (Batson, 2011; Batson, 2019; Capraro & Rand, 2018; Fehr & Fischbacher, 2003). For instance, Batson (2011; 2019) has shown that when people help out of *empathic concern*—an other-oriented emotion felt for someone in need—their actions reflect the ultimate goal of benefitting others (see also Eisenberg & Miller, 1987; Krebs, 1975; Hoffman, 1981). Psychological and neural evidence supporting empathy-based altruism continues to grow as well (Ashar et al., 2017; Batson, 2011; FeldmanHall et al., 2015; McAuliffe et al., 2018, Stocks et al., 2009; Weisz & Zaki, 2018).

Despite the strenuous scholarly debate about whether altruism exists or not, few studies have examined whether laypeople believe in altruism, and the consequences of such beliefs (Carlson & Zaki, 2018; Gebauer et al., 2015). As such, here we examine these beliefs and their psychological consequences.

Lay theories of altruism

Like scholars, laypeople have rich, intuitive theories about what drives people's behavior (Heider, 1958; Jara-Ettinger et al., 2016; Weiner, 1980), and such *lay theories* powerfully shape people's actions and inferences (Burnette et al., 2013; Hong, Levy, & Chiu, 2001; Ong et al., 2015; Schumann et al., 2014). For instance, people with a firmer belief in free will are less likely to cheat and steal (Vohs & Schooler, 2008; Vohs & Baumeister, 2009), and more likely to hold others culpable or punish them for moral transgressions (Baumeister & Brewer, 2012; Martin et al., 2017; Shariff et al., 2013).

Here we examined lay theories of altruism (Carlson & Zaki, 2018)—and specifically, peoples' beliefs about the role of *altruism* in their own and others' social behavior. Prior work suggests that such lay theories of altruism closely approximate what scholars term *psychological altruism* (Barasch et al., 2014; Carlson & Zaki, 2018; Swap, 1991)—which emphasizes an action's *motives*, as opposed to *biological altruism*, which emphasizes an action's reproductive fitness (Wilson, 2015). Thus, here we examined peoples' prior beliefs surrounding altruistic motivation: that is, beliefs about peoples' *capacity* for altruistic motives, and the *frequency* with which people act on altruistic motives. In particular, we sought to understand whether people's beliefs about altruism track with (i) how they perceive others' prosocial acts, and (ii) how often they behave prosocially themselves.

Effects on prosocial inferences. Lay theories of altruism could be an important guide for interpreting others' prosocial motives in daily life. Peoples' motives for acting prosocially are often ambiguous to observers. For instance, agents who act kindly usually reap benefits such as favors, praise, or good feelings. An onlooker could construe these benefits as the *reason* a person acted kindly—reflecting a self-oriented motive—or as an incidental *consequence* of their action (Batson, 2011)—consistent with a truly altruistic motive.

An observer's beliefs about altruism should “color in” this ambiguity: people who reject the existence of altruistic motives might tend to see others' kindness as self-oriented, whereas

people who believe in altruism might view it as genuine. However, in other cases, an agent's prosocial motives are *unambiguous*, such as when a friend confesses that they performed a kind act *in order to* impress a love interest. When an agent's true motives are explicitly known, an observer's lay theory of altruism should play little role in how they perceive the agent's prosocial act.

Effects on prosocial motivation. People's beliefs about altruism could also influence their *own* motivation to act prosocially (Gebauer et al., 2015). Those who reject the existence of altruistic motives should be more likely to believe that good deeds are performed *as a means* to achieving other goals—such as status or praise (i.e., *extrinsically* motivated), as opposed to *as ends* in themselves (i.e., *intrinsically* motivated, Kruglanski et al., 2018). But many studies suggest that associating prosocial acts with extrinsic motives, as opposed to intrinsic motives, *reduces* peoples' engagement in those actions (Batson et al., 1978; Heyman & Ariely, 2004; Hornstein, 1970; Kunda & Schwartz, 1983; Lin et al., 2017; Stukas et al., 1999; Uranowitz, 1975). This decline in prosociality may be due to at least three factors. First, associating prosocial acts with extrinsic (over intrinsic) motives can make one's actions more *conditional* on those extrinsic incentives (Lepper, Greene & Nisbett, 1973; Weinstein & Ryan, 2014). For example, if Jane primarily sees volunteering *as a means* to praise from other people, her volunteering may become more contingent on whether she sees a given opportunity as praiseworthy—thus narrowing the scope of volunteer opportunities that Jane will act upon. Second, associating prosocial acts with extrinsic over intrinsic motives can make prosociality more *expendable* in the pursuit of one's goals (Heider, 1958). If Jane believes that praise is the primary reason she volunteers, she should be more likely to cease volunteering if she finds a better way to be praised (e.g., academic achievement)—and thus, her interest in volunteering should be less stable over time. Lastly, associating prosocial acts with extrinsic relative to intrinsic motives can make prosociality seem *inauthentic* (Gebauer et al., 2015; Silver et al., 2021). If Jane sees volunteering as merely a means to an end, she should be less inclined to

view this activity as an honest and authentic way of expressing herself (e.g., her identity and values, Monin & Miller, 2016; Wiwad & Aknin, 2017)—further undermining her proclivity for this behavior.

In sum, extrinsic motives can be weaker sources of prosociality insofar as they make prosociality (i) *conditional* on the presence of extrinsic rewards, (ii) *dispensable* as a path to pursuing such extrinsic rewards, and (iii) *inauthentic* as a form of self-expression. By contrast, motives that are intrinsic to prosocial acts—such as the desire to benefit others (altruistic motivation)—should be stronger sources of prosociality precisely because these motives make prosocial acts *unconditional* and *indispensable* for satisfying one's goal (Batson, 2011), and because acting on such motives is viewed as *authentic* (Barasch et al., 2016).

Taken together, these findings suggest that lay theories of altruism could be self-fulfilling, with firmer beliefs in altruistic motives driving people to engage in *more* selfless behavior, and fainter beliefs driving people to engage in *less* selfless behavior.

Interactions with emotion and reason. What about a person could sway them towards believing or disbelieving in altruism? One key factor could be peoples' reliance on emotion versus reason when making prosocial judgments and decisions (Capraro, 2019; Levine et al., 2018). Recent work suggests that people consider empathy to be intimately tied to selfless behavior. For instance, donors who display stronger feelings for the victims they supported are perceived as more genuinely altruistic (Barasch et al., 2014). As such, individuals with a greater tendency to *feel* empathic concern, and thus to help others out of empathy (Batson, 2011), should also be more likely to believe in altruism (Gebauer et al., 2015).

On the other hand, individuals who tend to rely on reason might be less inclined to believe in altruism. Like Hobbes, laypeople are adept at generating self-oriented explanations for peoples' behavior (Miller, 1999), and the longer people spend reflecting on the motives behind their own or others' prosociality, the more self-oriented they adduce those motives to be (Critcher & Dunning, 2011; Batson et al., 1987). Together, this work suggests that people with a

greater tendency to engage in cognitive reflection in interpersonal contexts should be *less likely* to believe in altruistic motivation.

Self-other discrepancies. Beliefs about altruism could also shift depending on *whom* the beliefs apply to. People might believe *they themselves* are capable of altruism, and *other people* are as well (i.e., seeing oneself as an altruist among altruists). Alternatively, people might believe that they are *more* altruistic than others (i.e., seeing oneself as an altruist among egoists). Prior work suggests that people overestimate their generosity (Carlson, Maréchal, et al., 2020)—especially when comparing themselves to others (Epley & Dunning, 2000; Pronin, 2007; Rempel et al., 1985; but see also Klein & Epley, 2016). As such, beliefs about altruism may feature a motivated bias that places oneself *above* the average person (Kunda, 1990).

Another possibility is that people might believe that they are *less* altruistic than others. Nonmotivated biases can also loom large when comparing oneself to others (Chambers & Windschitl, 2004; Moore & Small, 2007)—and at times lead people to believe they are *below average*. Particularly, next to others, people tend to *underestimate* their ability to perform feats that they believe to be *difficult* (e.g., juggling; Kruger, 1999) or *uncommon* (e.g., seeing a comet in the sky; Chambers et al., 2003). Thus, to the extent that people believe altruism is a challenging or rare feat, they might tend to believe that they are even *less* capable of altruism than the average person.

To explore these different possibilities, it is important to assess both peoples' beliefs about their own and other peoples' *capacity* for altruism, as well as their beliefs about the *frequency* with which they and others engage in altruism.

Present research

In the present research, we examined whether lay theories of altruism predict the likelihood that people will both engage in good deeds themselves, and perceive others' good deeds as truly altruistic. Since initial work has focused on the positive relationship between belief in altruism and prosocial behavior (Gebauer et al., 2015), we first sought to replicate this

finding (Experiment 1). Next, we sought to extend this work by testing whether these beliefs positively track with peoples' prosocial inferences (Experiment 2)—which may ultimately serve as a 'bridge' between beliefs and behavior. In addition, we examined whether beliefs about altruism are related to peoples' disposition towards empathy and cognitive reflection. Lastly, we examined the internal consistency of these beliefs, comparing peoples' beliefs about their own versus others' propensity for altruism.

Experiment 1

In Experiment 1, we probed the relationship between belief in altruism and prosocial behavior. Specifically, we examined whether believing in altruism correlates with peoples' general prosocial tendencies, as well as their willingness to act prosocially in a laboratory task.

Method

We aimed for a minimum sample size of $N = 125$ in order to attain approximately 80% power to detect a medium-sized effect ($r \approx 0.25$; $\alpha = .05$). We recruited 150 participants from Amazon Mechanical Turk for an online experiment, and 150 completed the experiment. The experiment took approximately 10 minutes to complete.

This research was approved by the Ethical Committee of the Institutional Review Board at Stanford University. Written consent was obtained from all participants.

Belief in altruistic motives. To probe peoples' *belief in altruism*, participants completed a modified version of the Belief in the Existence of True Altruism scale (BETA; Gebauer et al., 2015; Gebauer et al., in prep). On this scale, people rated their agreement with 10 statements ($\alpha = .93$) about the nature of altruism, such as "Everything humans do for other people is motivated by self-interest of one kind or another". Ratings were made on a 7-point scale ranging from 1 ("Strongly disagree") to 7 ("Strongly agree").

Prosociality measures. We then assessed individuals' prosociality across three dimensions that were predicted to correlate with beliefs in altruism based on prior work

(Gebauer et al., 2015). First, we measured *trait prosocial tendencies* to assess whether beliefs in altruism relate to an individual's stable, prosocial habits. To this end, we asked participants to complete the Self-Report Altruism Scale (Rushton et al., 1981). This 20-item scale is designed to measure the frequency that one engages in prosocial acts primarily toward strangers. For instance, participants read "I have offered my seat on a bus or train to a stranger who was standing", and rated the frequency with which they've performed the act on a 5-point scale ranging from 1 ("Never") to 5 ("Very often").

Next, we measured *state prosocial behavior*. In a modified dictator game, all participants were told they would play the role of "decider" and "receiver". Participants learned that as deciders, they would receive an endowment, and decide how much of this money, if any, to transfer to the receiver (the next participant to complete the experiment). All participants received the same endowment size for their decision, independent of how much they received from another participant in their role as receiver. This endowment size (20 cents) was comparable to amounts used in prior work (Rand et al., 2012).

Participants read the following "How much money out of the 20 cents do you want to transfer to the next person completing this HIT (the Receiver)?" Participants then indicated the transfer amount on a slider scale (ranging from 0 – 20 cents in one-cent increments). After playing the role of decider, participants then played as a receiver, receiving any money transferred by the previous person who completed the experiment. Importantly, participants were told that their decisions as a decider would remain completely anonymous, and they would have no future interaction with other participants in the experiment after this task. This ensured that punishment or reputation could not be a strong motivating factor behind peoples' generosity.

Finally, we measured *hypothetical prosocial behavior*. Participants were presented with a hypothetical scenario, and were asked to treat it as real and respond as accurately as possible. The scenario read as follows: "Imagine that you find a \$20 bill on the sidewalk. No

people are in sight so you place the bill in your wallet.” Participants were then asked “How much of this \$20 would you be willing to anonymously donate to charity?”. Participants then indicated the donation amount on a slider scale (ranging from 0 - 20 dollars in one-dollar increments).

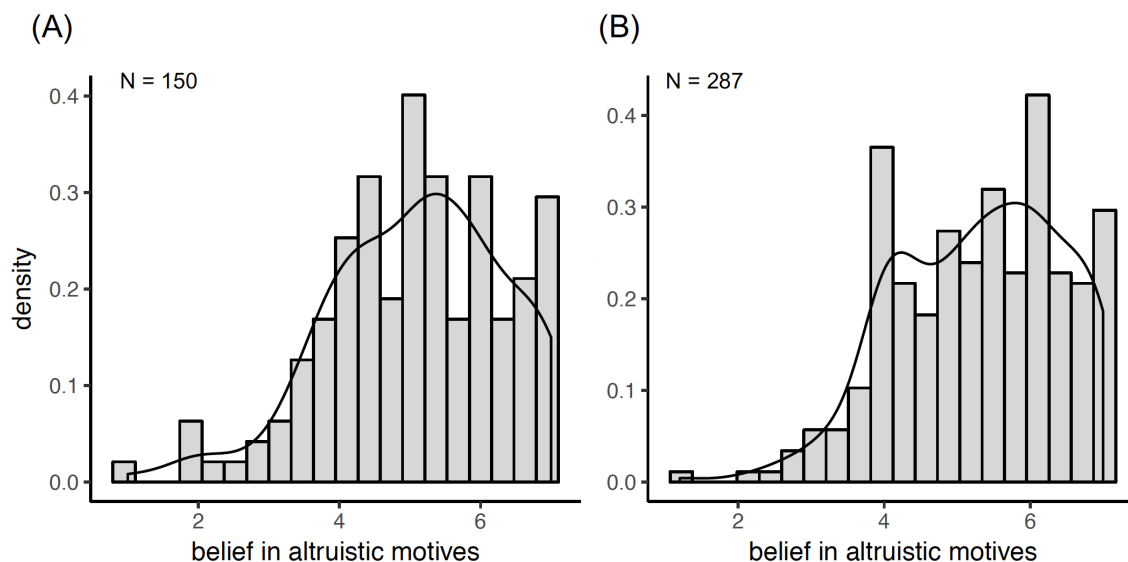
Full instructions for the dictator game and hypothetical donation can be found in the Supplemental Material (data and code can be found at <https://git.io/JtVJC>).

Results

People reported moderately strong beliefs in altruistic motives on average ($M_{BETA} = 5.10$, $SD_{BETA} = 1.25$, range = 1-7), consistent with preliminary work (Gebauer et al., 2015; see Figure 1a).

Figure 1

Distribution of beliefs in altruism (BETA) across experiment 1 (A), and experiment 2 (B). Lower values reflect stronger disbelief in altruism and higher values reflect stronger belief in altruism. We find that the majority of people hold a moderate-to-strong belief that humans can have altruistic motives.



Furthermore, the strength of peoples' belief in altruism positively correlated with all three of our measures of prosociality: trait prosocial tendencies ($r(148) = .21$, $p = .010$, 95% confidence interval (CI) [.05,.36]), hypothetical donations ($r(148) = .16$, $p = .045$, 95% CI [.004,.32]), and generosity in the dictator game ($r(148) = .21$, $p = .011$, 95% CI [.05,.36]). The

correlation between belief in altruism and hypothetical donations did not hold after controlling for peoples' trait prosocial tendencies ($r(147) = .11, p = .19, 95\% \text{ CI } [-.05, .27]$). However, the correlation between belief in altruism and generosity in the dictator game did hold even after controlling for trait prosocial tendencies ($r(147) = .18, p = .026, 95\% \text{ CI } [.02, .33]$) (See Table 1).

Table 1

Correlations between belief in altruism and three measures of prosociality. Belief in altruism correlates with generosity in a dictator game even after partialing out trait prosociality.

	<i>r</i>	<i>r</i> (TP partialled out)
trait prosociality (TP)	.21*	
hypothetical donating	.16*	.11
generosity	.21*	.18*

Note. * $p < .05$, ** $p < .005$, *** $p < .001$

Overall, these data support the idea that an individual's beliefs about altruism correlate with their prosociality, broadly speaking. Moreover, belief in altruism predicted contributions in the dictator game even after controlling for trait prosocial tendencies, suggesting that these beliefs could be a unique predictor of prosociality. In Experiment 2, we extended this finding by examining how these beliefs track not only people's own prosocial behavior, but their perceptions of others' prosociality as well.

Experiment 2

In Experiment 2, we tested whether peoples' beliefs about altruistic motives are related to how they perceive others' prosocial acts. Specifically, we predicted that belief in altruism would track perceptions of altruism when a prosocial agent's motives are *ambiguous* (i.e., agents received self-oriented benefits from their action), but not when an agent's motives are *unambiguous* (i.e. agents' actions were explicitly motivated by self-oriented benefits). Moreover, we examined whether these beliefs are positively associated with trait empathy and negatively associated with trait cognitive reflection. Furthermore, we examined peoples' beliefs about their

own, and others' propensity for altruism. Lastly, we sought to conceptually replicate Experiment 1 by showing that lay beliefs in altruism correlate with prosocial behavior, even when controlling for the influence of trait empathy.

Method

We assessed the relationship between lay theories of altruism and prosocial behavior within-subjects. Thus, we sought a minimum sample size of $N = 210$ in order to attain approximately 80% power to detect the smallest effects observed in Experiment 1 ($r \approx 0.19$; $\alpha = .05$). We assessed the relationship between belief's in altruism and perceptions of altruism between-subjects, where participants either learned of: (i) ambiguous motives, (ii) unambiguous motives, and (iii) no motive information (control condition). For such between-subject analyses, we aimed for $N = 90$ (per condition) in order to attain approximately 80% power to detect a medium-sized effect ($d \approx 0.30$; $\alpha = .05$). We recruited 300 participants from Amazon Mechanical Turk, and 295 completed the experiment. Participants who spent less than 1.5 s reading each prosocial scenario ($N = 8$) were excluded from analysis, as it was not possible to read the vignettes in this time span. This left a total sample size of $N = 287$ for our between-subject analyses (ambiguous motive condition = 88, unambiguous motive condition = 94, and control condition = 105). In addition, nine participants chose to not respond to our prosociality measure, leaving a sample size of $N = 278$ for these within-subject analyses. The experiment took approximately 12 minutes to complete.

This research was approved by the Ethical Committee of the Institutional Review Board at Stanford University. Written consent was obtained from all participants.

Prosocial scenarios. After consenting to participate, participants read about and assessed eight prosocial scenarios in an online experiment. Each scenario described an agent who saw and acted on an opportunity to help someone. For instance, in one such scenario participants read that Jane noticed a billboard urging passersby to donate blood, and decided to give blood at a local clinic (See Supplemental Material).

Crucially, participants were randomly assigned to one of three *within-subjects* conditions—which varied in how much information they provided about the prosocial agent's motives in each scenario. In the *ambiguous motive* condition, participants learned that the agent in each scenario may have acted on a self-oriented motive (i.e., the agent received a self-oriented benefit from their act). In the *unambiguous motive* condition, participants learned that the agent was explicitly motivated to obtain a self-oriented benefit from their act. In the *no motive information* condition, participants read scenarios only describing prosocial acts, with no mention of the agent's motives. This served as our control condition, providing a baseline for assessing the perceived altruism of a given action absent any other information.

Each scenario in the two experimental groups (the ambiguous motive and unambiguous motive conditions) included one further dimension—*how* the agent benefited. The self-oriented benefit in each scenario was either *material* (e.g. a tax break), *social* (e.g. praise), or *emotional* (e.g. feeling good)¹. These benefit types were counterbalanced across scenarios. We collapsed across benefit types to capture the more general effect of self-oriented benefits on inferences. However, our results hold for each of these benefit types when assessed independently (See Supplemental Table 1).

To illustrate, whereas participants in the no information condition would learn that Jane decided to give blood at a local clinic, those in the unambiguous motive condition might further read that Jane gave blood *in order* to receive a gift card (i.e., a material benefit), and those in the ambiguous motive condition might instead further read that *as a result* of giving blood, Jane impressed her friends (i.e., a social benefit). Crucially, while Jane benefits from her good deed in the latter scenario, it is left unclear in this condition whether the self-oriented benefit was her true motive—allowing a perceiver's own beliefs to potentially guide their inferences.

¹ Participants also read about a fourth type of benefit, but this benefit was other-oriented, and it was included to address a separate research question (see Carlson & Zaki, 2018). Since our focus was the effect that self-oriented benefits have on prosocial inferences, this fourth condition was not assessed here. Thus, though participants read 8 scenarios in the experimental conditions, we report analyses for only six.

Ratings of perceived altruism. After reading about each scenario, participants made a series of judgments about it. Specifically, participants rated (i) *how altruistic* they thought the prosocial *agent* was (ii) how altruistic they thought the agent's *action* was, and (iii) how altruistic they thought the agent's *motive* for their action was. For example, after reading about Jane giving blood, all groups were asked "How altruistic was Jane's action?", and responded on an 11-point scale ranging from 0 ("Not at all") to 10 ("Extremely"). Participants also rated *how selfish* they perceived agents, their actions, and their motives to be. For instance, after reading about Jane, participants were also asked: "How selfish is Jane as a person?" from 0 ("Not at all") to 10 ("Extremely").

Ratings for questions about agents, actions, and motives were highly correlated, $r = .86-.94$, $p < .001$. Thus, we collapsed these three ratings together, producing one composite measure of altruism ($\alpha = .98$), and one composite measure of selfishness ($\alpha = .96$).

Belief in altruistic motives. Next, to probe peoples' *belief in altruism*, participants again completed the BETA scale (Gebauer et al., 2015). Furthermore, to probe differences in lay theories of altruism with respect to the self versus others, participants responded to four additional, novel measures.

Specifically, we measured participants' beliefs about (i) their own *capacity* for altruism ("Do you believe that you are capable of pure altruism?"), as well as (ii) their own *frequency* of engaging in altruism ("When you perform a good deed, do you believe your motive is genuinely altruistic?"; both on a 5-point scale ranging from "Never" to "Always"). Moreover, we measured participants' beliefs about (iii) the average person's capacity for altruism ("Do you believe that the average person is capable of pure altruism?"), as well as (iv) the frequency with which the average person engages in altruism ("When the average person performs a good deed, do you believe their motive is genuinely altruistic?"; again using a 5-point scale ranging from "Never" to "Always").

Cognitive reflection. We then probed peoples' trait cognitive reflection to test whether those who should be relatively more prone to reflect on prosocial motives (high cognitive reflectors) are less likely to believe people act out of only altruistic motives (Critcher & Dunning, 2011). To this end, participants completed the Cognitive Reflection Test (CRT-2; Thomson & Oppenheimer, 2016).

Trait empathy. Next, to assess peoples' tendency toward other-oriented emotions, we measured two subcomponents of trait empathy—empathic concern and perspective taking. We focused on these two subcomponents of empathy because they tend to be related to prosociality (Davis, 2015). To this end, participants completed the empathic concern and perspective-taking subscales of the Interpersonal Reactivity Index (IRI; Davis, 1983).

Prosociality measure. Finally, to replicate the results of Experiment 1, we gave participants the opportunity to anonymously volunteer their time to complete up to two additional online tasks to help the experimenters. Participants chose whether they preferred to help with no additional tasks, a 2-minute task that involved generating recent life events, a 4-minute task that involved writing a profile of their life in general, or both tasks (6 minutes).

Results

People again reported moderately strong beliefs in altruism ($M_{BETA} = 5.30$, $SD_{BETA} = 1.14$, range = 1.2-7; Figure 1b).

In order to replicate our key finding from Experiment 1, we examined whether the strength of peoples' belief in altruism would predict their willingness to volunteer. Our results replicated Experiment 1, revealing that people with a stronger belief in altruism volunteered more time to complete an additional task ($r(276) = .18$, $p = .003$, 95% CI [.06,.29]). Moreover, this correlation held even when controlling for empathic concern and perspective taking ($r(274) = .16$, $p = .008$, 95% CI [.04,.27]) (See Table 2).

Table 2

Correlations between belief in altruism and prosociality after controlling for empathic traits. Belief in altruism correlates with volunteering even after partialing out empathic traits.

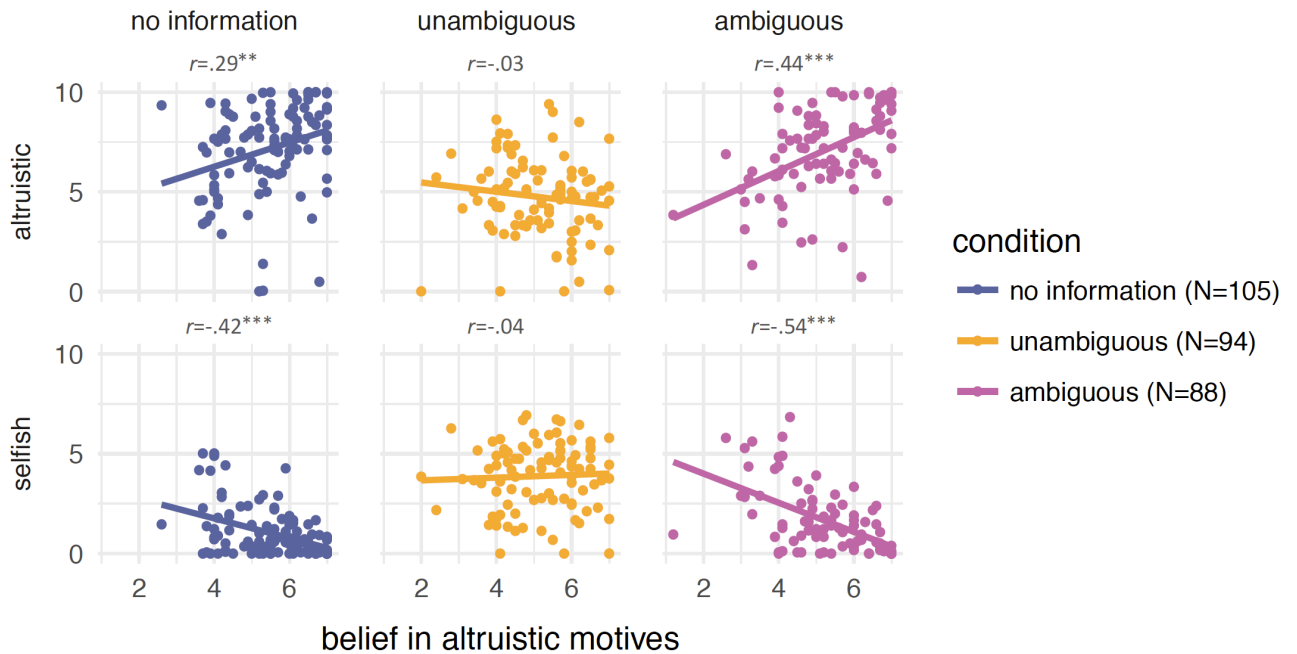
	<i>r</i>	<i>r</i> (EC/PT partialled out)
empathic concern (EC)	.33***	
perspective-taking (PT)	.28***	
volunteering	.18**	.16*

Note. * $p < .05$, ** $p < .005$, *** $p < .001$

Next, we examined whether peoples' lay theories of altruism shape the inferences they draw about others' good deeds. Specifically, we examined whether beliefs in altruism tracked with inferences that an agent had altruistic (versus selfish) intentions when people learned of (i) ambiguous motives (i.e., self-benefits resulted from the agents' good deed), (ii) unambiguous motives (i.e., self-benefits motivated the agents' good deed), or (iii) no motive-related information (control condition). We found that when a prosocial agent's motives were ambiguous, an observer's belief in altruism positively correlated with them rating prosocial agents as altruistically motivated ($r(86) = .48, p < .001, 95\% \text{ CI } [.30, .63]$), and negatively correlated with them rating prosocial agents as selfishly motivated ($r(86) = -.56, p < .001, 95\% \text{ CI } [-.69, -.39]$). However, as predicted, when a prosocial agent's motives were unambiguously selfish, observer's belief in altruism had no relationship with perceptions of altruism ($r(92) = -.13, p = .22, 95\% \text{ CI } [-.32, .08]$), or selfishness ($r(92) = .05, p = .67, 95\% \text{ CI } [-.16, .25]$). Finally, in the control condition, where observers learned no information about an agent's motives, belief in altruism again positively correlated with perceptions of altruism ($r(103) = .29, p = .003, 95\% \text{ CI } [.10, .45]$) and negatively correlated with perceptions of selfishness ($r(103) = -.42, p < .001, 95\% \text{ CI } [-.57, -.25]$) (See Figure 2). Means and standard deviations of ratings for each condition are reported in Table 3.

Figure 2

*Belief in altruism predicts inferences of altruism and selfishness behind prosocial acts when people learn no motive-related information (left panels) and when an actor's motives are ambiguous (right panels), but shows no relationship when an actor's motives are unambiguously self-interested (center panels). * $p < .05$, ** $p < .005$, *** $p < .001$.*

**Table 3**

Means and standard deviations for altruism and selfishness ratings for each condition.

condition	inference	M	SD
ambiguous motive	perceived altruism	7.12	2.20
	perceived selfishness	1.61	1.64
unambiguous motive	perceived altruism	4.74	2.02
	perceived selfishness	3.88	1.68
no information	perceived altruism	7.15	2.23
	perceived selfishness	1.03	1.24

Note. Ratings were made on an 11-point scale where: 0 = *Not at all*, 10 = *Extremely*

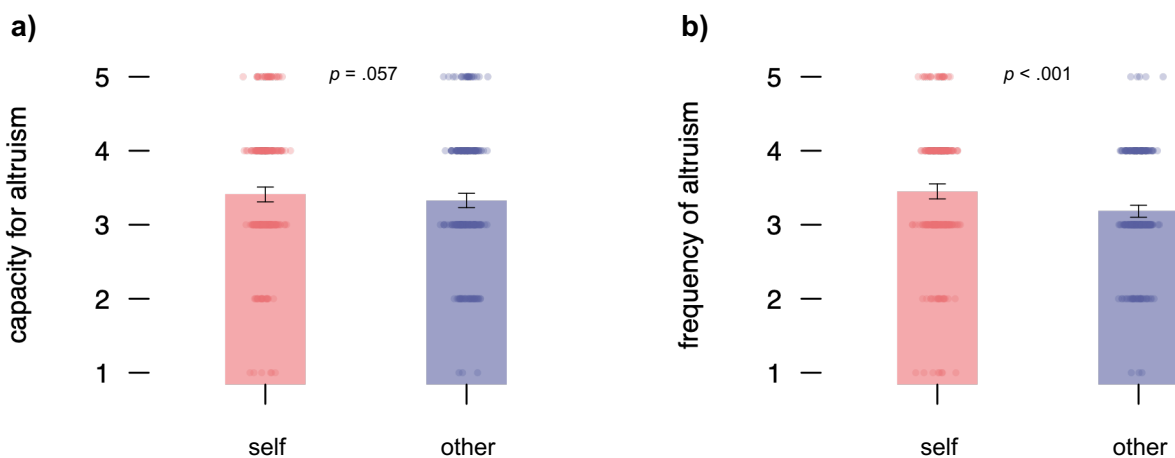
We also assessed the relationship between belief in altruism and our trait measures of empathy and cognitive reflection. Consistent with our predictions, belief in altruism positively correlated with empathic concern ($r(285) = .33$, $p < .001$, 95% CI [.23,.43]), and perspective-taking ($r(285) = .28$, $p < .001$, 95% CI [.17,.39]). However, contra our predictions, cognitive

reflection was not significantly negatively correlated with belief in altruism ($r(285) = .04$, $p = .52$, 95% CI $[-.08, .15]$).

Finally, we examined peoples' beliefs about altruism in themselves and others. Our results confirmed that people higher on our key measure of belief in altruism—the BETA scale—also more strongly believed that both they and others are capable of altruism (self: $r(285) = .38$, $p < .001$, 95% CI $[.28, .48]$; others: $r(285) = .41$, $p < .001$, 95% CI $[.30, .50]$). Moreover, people higher in belief in altruism also believed that they and others more frequently engage in altruism (self: $r(285) = .39$, $p < .001$, 95% CI $[.29, .49]$; others: $r(285) = .40$, $p < .001$, 95% CI $[.30, .50]$). We next directly compared peoples' beliefs about themselves versus others. People did not believe they were significantly more *capable* of altruism than others ($t(286) = .09$, $p = .057$, 95% CI $[-.003, .18]$, $d = .10$). However, people did believe they engage in altruism more *frequently* than others ($t(286) = .26$, $p < .001$, 95% CI $[.18, .35]$, $d = .33$; see Figure 3). Interestingly, belief in altruism had no relationship with the extent to which people believed they were more capable of altruism than others ($r(285) = -.02$, $p = .73$, 95% CI $[-.14, .10]$), or more frequently engaging in altruism than others ($r(285) = .06$, $p = .32$, 95% CI $[-.06, .17]$).

Figure 3

Beliefs about one's a) capacity for altruism and b) frequency of altruism as rated for self and other. We find no difference in beliefs about capacity), however we find that people believe they are significantly more altruistic than others). Error bars reflect 95% confidence intervals.



Discussion

Scholars have long debated the existence of altruistic motivation, but few have considered whether laypeople believe that altruism exists, or the implications of that belief. Across two experiments, we demonstrate that believing in altruism positively tracks with social behavior and inferences. People who believe in altruism not only perceived others' prosocial actions as more selfless, but also engaged in more selfless behavior themselves. Importantly, lay theories of altruism predicted peoples' generosity even after controlling for empathic concern and perspective-taking—two key predictors of prosocial behavior. These findings confirm prior work linking lay theories of altruism to moral character (Gebauer et al., 2015), and extend this work by clarifying *when* these beliefs guide inferences about others' prosocial motives.

Belief in altruism was highly correlated with perceptions of altruism when people learned that an agent benefited from performing their good deed, and thus the agent's motives were ambiguous. However, when people learned that an agent's motives for performing a good deed were unambiguously self-oriented, beliefs about altruism had no influence on how people perceived an agent's good deed. These findings are important for two reasons. First, they lend support to the idea that lay theories of altruism are used by observers specifically to predict a prosocial agent's *motives*, and play little to no role in assessing actions where one's motives are already known. Second, these findings suggest that lay theories of altruism do not merely reflect an optimism bias, as people endorsing stronger beliefs in altruism were just as cynical as those endorsing weaker beliefs when assessing good deeds that were motivated by self-interest. Furthermore, belief in altruism correlated with perceptions of altruism even when people did not learn of any benefits received by prosocial agents—though not as strongly as when motives were ambiguous. This suggests that in addition to guiding inferences when one must assess the relevance of a self-oriented benefit to an agent's *motives*, one's belief in altruism may also guide prosocial inferences in general.

We also found that people higher in empathic concern were more likely to endorse stronger beliefs in altruism. This finding offers indirect support for the idea that empathic concern may boost peoples' beliefs in altruism by increasing peoples' engagement in empathy-based altruism (Gebauer et al., 2015). However, contra our predictions, peoples' tendency toward cognitive reflection was unrelated to their beliefs about altruism. One possible reason for this finding is that the measure we used was not optimal for capturing peoples' tendency to engage in cognitive reflection in interpersonal situations. Indeed, prior work demonstrating a relationship between reflection and self-oriented attributions of behavior relied on reaction times (Critcher & Dunning, 2011) and explicit instruction to reflect (Batson et al., 1987). However, it is also possible that cognitive reflection might either weaken or dampen beliefs in altruism depending on the person and situation. This possibility aligns with recent work suggesting that the link between cognitive reflection and prosociality depends critically on the patterns of social interaction a person tends to encounter in daily life (Rand et al., 2014; Rand et al., 2016).

Lastly, we assessed peoples' beliefs about their own, and other peoples', tendencies towards altruism. In line with prior work on self-serving biases (Carlson, Maréchal, et al., 2020; Epley & Dunning, 2000), we found that people believe they engage in more altruistic behavior than others. Interestingly, people also believed they were only marginally more capable of altruism than others. In other words, people thought that while they have a similar capacity for altruism as others, their actions tend to be more altruistic. Interestingly, these effects had no relation to the strength of one's beliefs in altruism. Such a uniform bias in viewing one's self as more altruistic than others coalesces with prior work on self-enhancement (Brown, 1986b; Monin & Jordan, 2009), and aligns with the proposed importance of altruistic traits to peoples' self-concept (Aquino & Reed, 2002).

We opened this article by highlighting the Hobbesian view of human altruism—that *all* acts of kindness reflect veiled self-interest. However, taken together, our findings suggest that this view is decidedly rare among laypeople. Most people appear to believe that, at least

sometimes, people act out of an altruistic concern for the welfare of others. And crucially, the extent to which people believe that such altruism is possible, and prominent in human life, appears to be associated with them *seeing* more altruism, and *acting* more kindly themselves.

Limitations and future directions

In this work, we highlight a role for lay theories in social perception and behavior. However our research does not address how people form lay theories of altruism to begin with. As such, one important and exciting direction for future work to consider is how differences in peoples' beliefs about altruism develop (Cialdini et al., 1981; Perry et al., 1986), for instance through exposure to prosocial norms (Clark & Mills, 1979; Miller, 1999; Nook et al., 2016; Zaki et al., 2021), or popular science accounts of altruism and self-interest (Dawkins, 1976).

Additionally, in this work, we theorized that believing in altruistic motives tracks with prosociality, broadly construed. However, our experiments measured prosociality through actions that were largely anonymous and impersonal: namely, through anonymous donating and volunteering, as well as a trait measure of one's tendency to help strangers. As such, future work should examine a broader range of prosocial settings—especially those that vary in visibility and social distance to the target. Such work can help clarify the *breadth* of the connection between peoples' belief in altruistic motives, and their prosociality.

Lastly, while we demonstrate that belief in altruistic motives predicts prosocial actions and inferences, we cannot conclude whether these beliefs causally influenced these outcomes, or vice-versa. Thus, future work should expand on the present research by attempting to directly *manipulate* peoples' beliefs about altruistic motives. Such work would no doubt deepen our understanding of the moral “force” that lay theories of altruism carry in human life.

References

- Aquino, K., & Reed, I. I. (2002). The self-importance of moral identity. *Journal of personality and social psychology*, 83(6), 1423.
- Andreoni, J., & Rao, J. M. (2011). The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of Public Economics*, 95(7), 513-520.
- Ashar, Y. K., Andrews-Hanna, J. R., Dimidjian, S., & Wager, T. D. (2017). Empathic care and distress: Predictive brain markers and dissociable brain systems. *Neuron*, 94(6), 1263-1273.
- Barasch, A., Berman, J. Z., & Small, D. A. (2016). When payment undermines the pitch: On the persuasiveness of pure motives in fund-raising. *Psychological Science*, 27(10), 1388-1397.
- 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(3), 393.
- Batson, C. D. (2011). *Altruism in humans*. Oxford University Press.
- Batson, C. D. (2019). *A scientific search for altruism: Do we care only about ourselves?*. Oxford University Press.
- Batson, C. D., Coke, J. S., Jasnosi, M. L., & Hanson, M. (1978). Buying kindness: Effect of an extrinsic incentive for helping on perceived altruism. *Personality and Social Psychology Bulletin*, 4(1), 86-91.
- Batson, C. D., Fultz, J., Schoenrade, P. A., & Paduano, A. (1987). Critical self-reflection and self-perceived altruism: When self-reward fails. *Journal of personality and social psychology*, 53(3), 594.
- Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. *The American Economic Review*, 97(2), 170-176.
- Baumeister, R. F., & Brewer, L. E. (2012). Believing versus disbelieving in free will: Correlates and consequences. *Social and Personality Psychology Compass*, 6(10), 736-745.
- Bodner, R., & Prelec, D. (2003). Self-signaling and diagnostic utility in everyday decision making. *The Psychology of Economic Decisions*, 1, 105-26.
- Brown, J. D. (1986a). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, 4(4), 353-376.
- Brown, R. (1986b). Altruism and affection. In R. Brown (2nd ed.), *Social psychology*, (pp. 89-124). Free Press.
- Burnette, J. L., O'boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychological bulletin*, 139(3), 655.

- Capraro, V. (2019). The dual-process approach to human sociality: A review. *Available at SSRN 3409146*.
- Capraro, V., & Rand, D. G. (2018). Do the right thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality. *Judgment and Decision Making*, 13(1), 99-111.
- Carlson, R. W., Adkins, C., Crockett, M. J., & Clark, M. S. (2020). Psychological selfishness. <https://doi.org/10.31234/osf.io/jp27m>
- Carlson, R. W., Maréchal, M. A., Oud, B., Fehr, E., & Crockett, M. J. (2020). Motivated misremembering of selfish decisions. *Nature Communications*, 11(1), 1-11.
- Carlson, R. W., & Zaki, J. (2018). Good deeds gone bad: Lay theories of altruism and selfishness. *Journal of Experimental Social Psychology*, 75, 36-40.
- Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of personality and social psychology*, 37(1), 12.
- Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: the role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin*, 130(5), 813.
- Chambers, J. R., Windschitl, P. D., & Suls, J. (2003). Egocentrism, event frequency, and comparative optimism: When what happens frequently is “more likely to happen to me”. *Personality and Social Psychology Bulletin*, 29(11), 1343-1356.
- Cialdini, R. B., Baumann, D. J., & Kenrick, D. T. (1981). Insights from sadness: A three-step model of the development of altruism as hedonism. *Developmental Review*, 1(3), 207-223.
- Critcher, C. R., & Dunning, D. (2011). No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology*, 47(6), 1207-1213.
- Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67-80.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113-126.
- Davis, M. H. (2015). Empathy and Prosocial Behavior. *The Oxford Handbook of Prosocial Behavior*, 282.
- Dawkins, R. (1976). *The selfish gene*. Oxford University Press.
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, 101(1), 91.

- Epley, N., & Dunning, D. (2000). Feeling "holier than thou": are self-serving assessments produced by errors in self-or social prediction?. *Journal of Personality and Social Psychology*, 79(6), 861.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785.
- FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage*, 105, 347-356.
- Freud, S. (1930). *Civilization and its discontents*.
- Gebauer, J. E., Sedikides, C., Leary, M. R., & Asendorpf, J. B. (2015). Lay beliefs in true altruism versus universal egoism. *Character: New Directions from Philosophy, Psychology, and Theology*, 75.
- Gebauer, J. E., Sedikides, C., Leary, M. R., & Asendorpf, J. B. (in prep). Belief in the existence of true altruism (BETA) scale.
- Heider, F. (1958). *The psychology of interpersonal relations*. Psychology Press.
- Heyman, J., & Ariely, D. (2004). Effort for payment: A tale of two markets. *Psychological Science*, 15(11), 787-793.
- Hobbes, T. (1651). *Leviathan: Or the matter, form, and power of a commonwealth, ecclesiastical and civil*.
- Hoffman, M. L. (1981). Is altruism part of human nature? *Journal of Personality and Social Psychology*, 40, 121-137.
- Hong, Y. Y., Levy, S. R., & Chiu, C. Y. (2001). The contribution of the lay theories approach to the study of groups. *Personality and Social Psychology Review*, 5(2), 98-106.
- Hornstein, H. A. (1970). The influence of social models on helping. In Macaulay, J., & Berkowitz, L. (Eds.), *Altruism and helping behavior: social psychological studies of some antecedents and consequences* (pp. 29-41). New York, NY: Academic Press.
- Hume, D. (1902). *An inquiry concerning the principles of morals* (L. A. Selby-Bigge, Ed.). Oxford: Oxford University Press (first published, 1751).
- Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, 20(8), 589-604.
- Klein, N., & Epley, N. (2016). Maybe holier, but definitely less evil, than you: Bounded self-righteousness in social judgment. *Journal of Personality and Social Psychology*, 110(5), 660.
- Krebs, D. (1975). Empathy and altruism. *Journal of Personality and Social Psychology*, 32(6), 1134.

- Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77(2), 221.
- Kruglanski, A. W., Fishbach, A., Woolley, K., Bélanger, J. J., Chernikova, M., Molinario, E., & Pierro, A. (2018). A structural model of intrinsic motivation: On the psychology of means-ends fusion. *Psychological Review*, 125(2), 165.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, 108(3), 480.
- Kunda, Z., & Schwartz, S. H. (1983). Undermining intrinsic moral motivation: External reward and self-presentation. *Journal of Personality and Social Psychology*, 45(4), 763.
- Lenton, A. P., Bruder, M., Slabu, L., & Sedikides, C. (2013). How does "being real" feel? The experience of state authenticity. *Journal of Personality*, 81(3), 276-289.
- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129.
- Levine, E. E., Barasch, A., Rand, D., Berman, J. Z., & Small, D. A. (2018). Signaling emotion and reason in cooperation. *Journal of Experimental Psychology: General*, 147(5), 702.
- Lin, S. C., Zlatev, J. J., & Miller, D. T. (2017). Moral traps: When self-serving attributions backfire in prosocial behavior. *Journal of Experimental Social Psychology*, 70, 198-203.
- Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The effects of perspective taking on motivations for helping: Still no evidence for altruism. *Personality and Social Psychology Bulletin*, 28(11), 1601-1610.
- Mansbridge, J. J. (1990). *Beyond self-interest*. University of Chicago Press.
- Martin, N. D., Rigoni, D., & Vohs, K. D. (2017). Free will beliefs predict attitudes toward unethical behavior and criminal punishment. *Proceedings of the National Academy of Sciences*, 114(28), 7325-7330.
- McAuliffe, W. H., Forster, D. E., Philippe, J., & McCullough, M. E. (2018). Digital altruists: Resolving key questions about the empathy-altruism hypothesis in an Internet sample. *Emotion*, 18(4), 493.
- Miller, D. T. (1999). The norm of self-interest. *American Psychologist*, 54(12), 1053.
- Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. *Personality, identity, and character: Explorations in moral psychology*, 341-354.
- Monin, B., & Miller, D. T. (2016). Moral opportunities versus moral tests. In *The social psychology of morality* (pp. 56-71). Routledge.
- Moore, D. A., & Small, D. A. (2007). Error and bias in comparative judgment: on being both better and worse than we think we are. *Journal of personality and social psychology*, 92(6), 972.

- Nook, E. C., Ong, D. C., Morelli, S. A., Mitchell, J. P., & Zaki, J. (2016). Prosocial conformity: Prosocial norms generalize across behavior and empathy. *Personality and Social Psychology Bulletin*, 42(8), 1045-1062.
- Ong, D. C., Zaki, J., & Goodman, N. D. (2015). Affective cognition: Exploring lay theories of emotion. *Cognition*, 143, 141-162.
- Perry, L. C., Perry, D. G., & Weiss, R. J. (1986). Age differences in children's beliefs about whether altruism makes the actor feel good. *Social Cognition*, 4(3), 263-269.
- Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, 11(1), 37-43.
- Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416), 427.
- Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95.
- Rousseau, J. J. (1755). *The social contract and discourses*.
- Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, 2(4), 293-302.
- Schumann, K., Zaki, J., & Dweck, C. S. (2014). Addressing the empathy deficit: beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *Journal of Personality and Social Psychology*, 107(3), 475.
- Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. J., Schooler, J. W., ... & Vohs, K. D. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science*, 25(8), 1563-1570.
- Silver, I., Newman, G., & Small, D. A. (2021). Inauthenticity aversion: Moral reactance toward tainted actors, actions, and objects. *Consumer Psychology Review*, 4(1), 70-82.
- Skinner, B. F. (1978). The ethics of helping people. In L. Wispé (Ed.), *Altruism, sympathy, and helping: Psychological and sociological principles* (pp. 249-262). New York: Academic Press.
- Smith, A. (1759). *The theory of moral sentiments*.
- Stocks, E. L., Lishner, D. A., & Decker, S. K. (2009). Altruism or psychological escape: Why does empathy promote prosocial behavior?. *European Journal of Social Psychology*, 39(5), 649-665.
- Stukas, A. A., Snyder, M., & Clary, E. G. (1999). The effects of "mandatory volunteerism" on intentions to volunteer. *Psychological Science*, 10(1), 59-64.
- Swap, W. C. (1991). When prosocial behavior becomes altruistic: An attributional analysis. *Current Psychology*, 10(1), 49-64.

- Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1), 99-113.
- Uranowitz, S. W. (1975). Helping and self-attributions: A field experiment. *Journal of Personality and Social Psychology*, 31(5), 852.
- Vohs, K. D., & Baumeister, R. F. (2009). Addiction and free will. *Addiction Research & Theory*, 17(3), 231-235.
- Vohs, K. D., & Schooler, J. W. (2008). The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science*, 19(1), 49-54.
- Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology*, 39(2), 186-200.
- 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.
- Weisz, E. & Zaki, J. (2018). Motivated empathy: A social neuroscience perspective. *Current Opinion in Psychology*.
- Wilson, D. S. (2015). *Does altruism exist?: culture, genes, and the welfare of others*. Yale University Press.
- Wiwad, D., & Aknin, L. B. (2017). Motives matter: The emotional consequences of recalled self- and other-focused prosocial acts. *Motivation and Emotion*, 41(6), 730-740.
- Zaki, J., Neumann, E., & Baltiansky, D. (2021). Market cognition: How exchange norms alter social experience. <https://doi.org/10.31234/osf.io/m5swc>