Belief in altruistic motives predicts prosocial actions and inferences

Ryan W. Carlson[1] & Jamil Zaki[2]

[1]Department of Psychology, Yale University

[2]Department of Psychology, Stanford University

Address correspondence to:

Ryan W. Carlson
Department of Psychology, Yale University
2 Hillhouse Avenue
New Haven, CT 06511
ryan.carlson@yale.edu

Abstract word count: 151
Main text word count: 4675

Abstract

Are humans ever truly altruistic? Or are all actions, however noble, ultimately motivated by self-interest? Scientists 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. In line with prior work, we find that people tend to believe humans can be, and frequently are, altruistically motivated. Moreover, people who more strongly believe in altruistic motivation act more prosocially themselves—for instance, sacrificing relatively high levels of money and time to help others—a relationship that holds even when controlling for trait empathy. People who believe in altruism also judge other prosocial agents to be more genuinely kind, especially when agents' motives are ambiguous. Together, this work suggests that believing in altruism predicts the extent to which people both *see* altruism and *act* altruistically, possibly reflecting the self-fulfilling nature of such lay theories.

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

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 (Andreoni & Rao, 2011; Battigalli and Dufwenberg, 2007; Bodner and Prelec, 2003; Dana et al., 2007; Maner et al., 2002).

Other philosophers argue that altruism does exist (Hume, 1902; Rousseau, 1754; Smith, 1759), and psychologists have also provided compelling evidence to back this view. 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).

Lay theories of altruism

Despite the strenuous scholarly debate about whether altruism exists or not, few studies have examined what laypeople believe about altruism and the consequences of

such beliefs (Carlson & Zaki, 2018; Gebauer et al., 2015). Like scholars, laypeople have rich, intuitive theories about what drives people's behavior (Heider, 1958; 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*—or the structured beliefs people hold about their own and others' prosocial motives (Carlson & Zaki, 2018). Prior work suggests that 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 *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 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 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 motives can make prosociality seem *inauthentic* (Gebauer et al., 2015). 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.

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 (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 at 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 selflessness (Carlson et al., 2018)—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 (Study 1). Next, we sought to extend this work by testing whether these beliefs positively track with peoples' prosocial inferences (Study 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.

Study 1

In Study 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 study, and received 150 completed surveys.

Lay theories of altruism. To probe peoples' *belief in altruism*, participants completed 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 stable, *trait prosocial tendencies*. 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.

Next, we measured *real prosocial behavior*. In a modified dictator game, participants were told they would play the role of "decider" and "receiver". Participants learned that as deciders, they would receive 20 cents, and decide how much of this money, if any, to transfer to the receiver (the next participant to complete the survey). Specifically, 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 survey. 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 study 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 response 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://github.com/carlsonrw/belief_altMotives)

Results

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

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% 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% CI [.02,.33]) (See Table 1).

Overall, these data support the idea that an individual's beliefs about altruism correlate with his or her prosociality, broadly speaking (Gebauer at al., 2015). 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 Study 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.

Study 2

In Study 2, we tested whether peoples' beliefs about altruism predict 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 personally benefit from their action), but not when an agent's motives are *unambiguous* (i.e. agents' actions were explicitly motivated by personal 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 Study 1 by showing that lay beliefs in altruism correlate with real 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 Study 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. 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 received 295 completed surveys. Nine participants were excluded from analysis for failing to follow task instructions, leaving 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.

Prosocial vignettes. Participants read and rated eight vignettes in an online survey. Each vignette described a prosocial action performed by an agent. For instance, in one such vignette participants read about Jane, who gave blood at a local clinic (See Supplemental Material).

Participants were randomly assigned to read vignettes in one of three conditions. Participants in the *control* condition read vignettes that only described prosocial actions. This provided a baseline for assessing the perceived altruism of a given action absent any other information. Participants in the other two experimental groups read about one of three types of benefits the agent received as a result of her prosocial action: (i) *material* benefits, for instance receiving a tax break after a charitable donation, (ii) *social* benefits, such as receiving praise for a donation, or (iii) *emotional* benefits, such as feeling good after making a donation. Participants read about a fourth (non-selfish) type of benefit (see Carlson & Zaki, 2018), which was not assessed in the present study. We collapsed across the three benefit types described above, as such comparisons are made elsewhere (Carlson & Zaki, 2018). However, our results hold across the three benefit types (See Supplemental Table 1).

Participants who read about prosocial agents benefitting from their actions were further randomized to read about these benefits either as *motivating* those actions (i.e., *unambiguous motive* condition) or as an incidental *consequence* of those actions (i.e., *ambiguous motive* condition). For instance, participants in the unambiguous motive group might read that Jane gave blood *in order* to (i) receive a gift card, (ii) impress her friends, or (iii) feel good. Participants in the ambiguous motive group might instead read that *as a result* of giving blood, Jane (i) received a gift card, (ii) impressed her friends, or (iii) felt good. Thus, while Jane benefitted from her good deed, it is left unclear in this condition whether this benefit was her true motive. Again, vignettes were counterbalanced, such that each participant read about each benefit type paired with each of two prosocial

actions. Unlike benefit type, ambiguous motive versus unambiguous motive versus control conditions were manipulated between subjects.

Ratings of perceived altruism. After reading each vignette, 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 and one composite measure of selfishness.

Prosocial opportunity.

In order to replicate the results of Study 1, we gave participants the opportunity to anonymously volunteer their time to complete up to two additional online tasks to help the researchers. 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).

Lay theories of altruism. To probe peoples' *belief in altruism*, participants again completed the BETA scale (Gebauer at 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-poit scale ranging from "Never" to "Always".

Trait empathy and cognitive reflection. We measured two subcomponents of trait empathy: empathic concern and perspective taking. Participants completed the empathic concern and perspective-taking subscales of the Interpersonal Reactivity Index (IRI; Davis, 1983). In addition, we probed peoples' trait cognitive reflection. To this end, participants completed the Cognitive Reflection Test (CRT-2; Thomson & Oppenheimer, 2016).

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 Study 1, we examined whether the strength of peoples' belief in altruism would predict their willingness to volunteer. Our results replicated Study 1, revealing that people with a stronger in 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).

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 ($r(86) = .44$, $p < .001$, 95% CI [.25,.59]), and negatively correlated with them rating prosocial agents as selfishly motivated ($r(86) = -.54$, $p < .001$, 95% CI [-.68,-.38]). However, as predicted, when a prosocial agent's motives were unambiguously selfish, observer's belief in altruism had no relationship with perceptions of altruism ($r(92) = -.03$, $p = .75$, 95% CI [-.23,.17]), or selfishness ($r(92) = -.04$, $p = .72$, 95% CI [-.24,.17]). 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% CI [.10,.45]) and negatively correlated with perceptions of selfishness ($r(103) = -.42$, $p < .001$, 95% CI [-.57,-.25]) (See Table 3).

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 considerably more *frequently* than others ($t(286) = .26$, $p < .001$, 95% CI [.18,.35], d = .33). Interestingly, belief in altruism had no relationship with the extent to which people believed themselves to be 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]).

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 studies, we demonstrate that believing in altruism predicts 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 at 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 his or her 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 good deeds. 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.

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 at al., 2015). However, contra our predictions, peoples' tendency toward cognitive reflection was unrelated to their beliefs

about altruism. One possible reason for this null 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).

Lastly, we assessed peoples' beliefs about their own, and other peoples', tendencies towards altruism. In line with prior work on self-serving biases (Carlson et al., 2018; Epley & Dunning, 2000), we found that people believed they were more consistently altruistic (but only marginally more capable of altruism) than others. Surprisingly, this effect 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).

Future directions

Our work highlights a role for lay theories in moral perception and behavior. Future work should also consider how differences in peoples' beliefs about altruism develop (Cialdini et al., 1981; Perry et al., 1986), for instance through exposure to prosocial norms (Miller, 1999; Nook et al., 2016) or popular science accounts of altruism (Dawkins, 1976). Future work could also expand on the present research by directly *manipulating* peoples' beliefs about altruism, either through interventions that target social norms surrounding altruism, or through selectively exposing people to scientific evidence *for* or *against* altruism. Such research could provide more direct evidence of the moral "force" that lay theories of altruism carry in human life.

References

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*.

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., Jasnoski, 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. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, *4*(4), 353-376.

Brown, R. (1986). 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), 655-701.

Carlson, R. W., & Zaki, J. (2018). Good deeds gone bad: Lay theories of altruism and selfishness. *Journal of Experimental Social Psychology*, *75*, 36-40.

Carlson, R. W., Marechal, M., Oud, B., Fehr, E., & Crockett, M. (2018). Motivated misremembering: Selfish decisions are more generous in hindsight. https://doi.org/10.31234/osf.io/7ck25

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.

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.

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).

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. B., 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), 493-506

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.

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.

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.

Table 1

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

	r	*r* (TP partialed out)
trait prosociality (TP)	.21*	
hypothetical donating	.16*	.11
generosity	.21*	.18*

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

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.

	r	*r* (EC/PT partialed out)
empathic concern (EC)	.33***	
perspective-taking (PT)	.28***	
volunteering	.18**	.16*

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

Table 3

Relationship between belief in altruism and perceiving altruism in other's good deeds when people learned of (i) ambiguous motives, (ii) unambiguous motives, or (iii) no motive-related information (control condition).

condition	inference	*r*
ambiguous motive	perceived altruism	.44***
N=88	perceived selfishness	-.54***
unambiguous motive	perceived altruism	-.03
N=94	perceived selfishness	-.04
no information	perceived altruism	.29**
N=105	perceived selfishness	-.42***

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

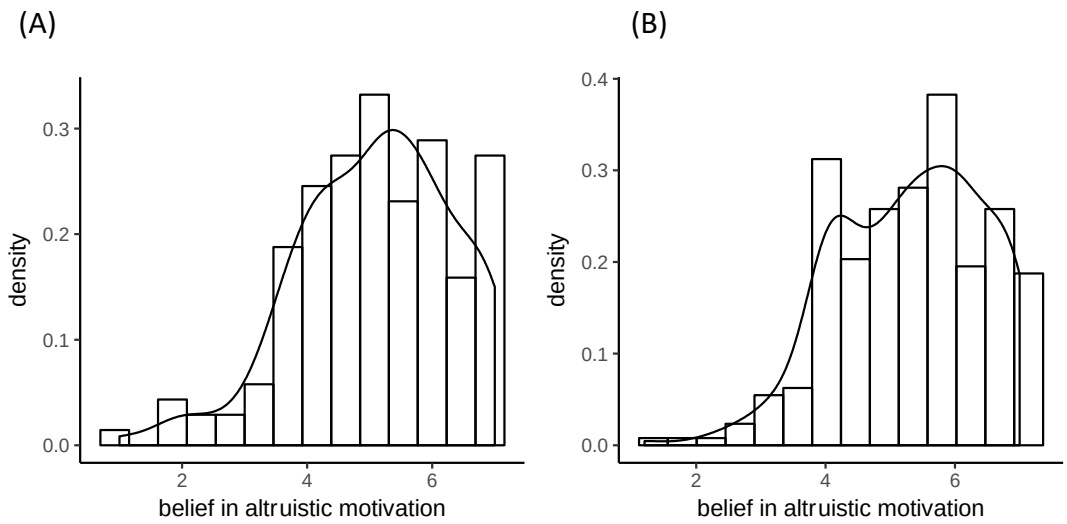

Figure 1. Here we show the distribution of beliefs in altruism (BETA) across study 1 (A), and study 2 (B). Lower values reflect stronger *disbelief* in altruism and higher values reflect stronger *belief* in altruism. In line with prior work (Gebauer et al., 2015), we find that the majority of people hold a moderate-to-strong belief in altruism.