Journal of Personality and Social Psychology

Anticipated Affect Predicts Moral Praise and Character Judgments

Paige Amormino, Brett Mercier, and Yoel Inbar

Online First Publication, January 11, 2024. https://dx.doi.org/10.1037/pspa0000377

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Personality and Social Psychology:
Attitudes and Social Cognition

ISSN: 0022-3514

https://doi.org/10.1037/pspa0000377

Anticipated Affect Predicts Moral Praise and Character Judgments

Paige Amormino[1], Brett Mercier[2], and Yoel Inbar[2]
[1] Department of Psychology, Georgetown University
[2] Department of Psychology, University of Toronto

In four preregistered studies (total $N = 5{,}067$), we investigated whether people use their own anticipated affective responses to a situation to make judgments about the praiseworthiness of helping and the moral character of helpers. We found that helpers in more affectively arousing scenarios were seen as more morally motivated, received greater praise, and were seen as having more positive moral character, even when controlling for the perceived benefits of helping. Describing helpers as unemotional reduced the effect of observers' anticipated affect on character judgments. These results suggest that when making praise and character judgments, people not only consider the consequences of a helper's actions but also the emotions they experience when engaging in them—and that they use their own anticipated emotional experience to do this.

Keywords: moral judgment, empathy, person perception, helping, prosociality

Supplemental materials: https://doi.org/10.1037/pspa0000377.supp

Be it opening the door for a stranger, offering up a seat on the subway, or rounding up purchases for charity, praiseworthy acts are ubiquitous. In daily life, people report committing more moral than immoral acts, being the target of more moral than immoral acts, and directly witnessing more moral than immoral acts (Hofmann et al., 2014). Compared to moral blame, judgments of moral praise are less well understood (Anderson et al., 2020), but it is clear that praise judgments are not simply the inverse of blame judgments, and that in some cases they rely on distinct psychological mechanisms (see, e.g., Guglielmo & Malle, 2019; Pizarro et al., 2003; Siegel et al., 2017).

The most important difference between praise and blame judgments may be that praise judgments are more sensitive to observer beliefs about the actor's underlying moral character (Anderson et al., 2020). When making moral judgments, people often act as intuitive virtue ethicists who care most about what the behavior signals about "what kind of person" the actor is (Carlson et al., 2022; Uhlmann et al., 2015). This particularly seems to be the case for judgments of moral praise. After all, someone who acts prosocially (e.g., donates a considerable sum to charity) may have done so for morally praiseworthy reasons (e.g., genuine concern for the needy) or less praiseworthy ones (e.g., demonstrating his generosity to observers).

People see the helper's emotional state as especially informative about whether they acted for the right reasons. When people consider hypothetical others who decided to help them based on emotion or cost–benefit analysis, they attribute more praiseworthy mental states to the emotionally motivated helpers (Ames et al., 2004). Likewise, people see emotionally motivated charity donors as more moral because they infer that emotion signals genuine moral concern for the recipients of aid (Barasch et al., 2014). People who donate to charity based on emotion rather than deliberation are seen more positively, again because they are seen as more morally motivated (Montealegre et al., 2021). Finally, people think that others ought to donate to causes to which they have an emotional connection, even when these are less effective at helping recipients (Berman et al., 2018).

Overall, then, prototypical prosocial behavior is seen as emotionally motivated, and emotionally motivated helpers are evaluated more positively (Barasch et al., 2014; Berman et al., 2018). But how do people infer that helping is emotionally motivated when they make real-world judgments? The research described above cannot answer this question because in these studies, participants are usually informed explicitly whether an agent's behavior was emotionally motivated or not (e.g., Ames et al., 2004; Barasch et al., 2014; Montealegre et al., 2021). This provides a high degree of experimental control, but it is not particularly realistic. It makes salient and explicit many aspects of praiseworthy actions and actors (i.e., their emotional state and internal motivations) that are, in everyday life, not presented to us but inferred. Thus, the goal of the current research is to understand how inferences about a helper's emotions are made, and how these affect people's evaluations of prosocial behavior and those who engage in it.

One suggestive area of research examines how moral evaluations of acts respond to "action value" (how good or bad an action would feel to perform) versus "outcome value" (how much an action helps or harms a target). Most of this research has examined harmful actions, for which wrongness judgments seem to respond equally to

Paige Amormino https://orcid.org/0000-0003-2125-1047
Brett Mercier https://orcid.org/0000-0003-3731-8752
Yoel Inbar https://orcid.org/0000-0002-3176-3727

This work was funded in part by Canadian Studies, Princeton University, awarded to Paige Amormino.

Paige Amormino played a lead role in conceptualization and data curation and an equal role in investigation, project administration, writing–original draft, and writing–review and editing. Brett Mercier played a lead role in formal analysis and visualization and a supporting role in data curation. Yoel Inbar played a lead role in funding acquisition, methodology, project administration, resources, and writing–review and editing, a supporting role in data curation, and an equal role in conceptualization, formal analysis, and writing–original draft.

Correspondence concerning this article should be addressed to Paige Amormino, Department of Psychology, Georgetown University, 3700 O Street NW, Washington, DC 20057, United States. Email: pa546@georgetown.edu

action value and outcome value (Cushman, 2008; Cushman et al., 2012; R. Miller & Cushman, 2013; R. M. Miller et al., 2014). For example, ratings of the moral wrongness of mercy killing carried out in different ways (e.g., shooting the victim with a gun, stabbing him in the throat, or suffocating him with a pillow) are predicted by both how much people think the victim would suffer in the act (i.e., outcome value) and by how upset they would be by harmlessly acting out the behavior as part of a movie plot (i.e., action value; R. M. Miller et al., 2014, Study 4). The single investigation of outcome and action value for prosocial behavior (Yudkin et al., 2019) found judgments of praise are likewise predicted both by how much people think that an action will benefit the person helped (i.e., outcome value) and by how good it would feel to act out the behavior as part of a movie plot (i.e., action value). It is unclear why there is a relationship between action value and praise judgments, though the authors speculated that such relationship might arise from people mentally simulating themselves in the helper's position. Helping typically feels good (i.e., the "warm glow"; Andreoni, 1990), and so the positive affect resulting from this mental simulation might affect intuitive evaluations of the helper's behavior (Yudkin et al., 2019; see also R. M. Miller et al., 2014).

In the current research, we build on the insight that, when judging prosocial actors, people may mentally simulate what it is like to be in the actor's place. However, we differ from past theorizing by proposing that the mental processes following this mental simulation may be more sophisticated than a simple association between positive affect and behavioral evaluations. Specifically, we hypothesize that people may use their own anticipated emotional reactions in a situation (i.e., their anticipated affect) to infer how other people would have felt in the same situation—and that these inferences may have downstream effects on moral judgment. We base this prediction on substantial literature showing that people often use their own knowledge, beliefs, and feelings to make judgments of others.

Social Projection and Inferred Motives for Helping

Social projection—the process of using information about the self to make judgments about others—is ubiquitous. Until they are given a reason to think otherwise, people assume others share their attitudes, beliefs, and behaviors (Krueger, 1998, 2000). When they need to make a judgment about whether other people would do something (e.g., whether they would agree to wear a large sandwich board displaying the slogan "Repent" in public), they start by asking whether they themselves would do it (Ross et al., 1977). Similarly, when people make predictions about the emotions other people would experience in a situation, they start by imagining how they themselves would feel—though they may subsequently adjust their judgments away from their egocentric starting point if they have reason to believe that their own experience is uninformative about that of others (Van Boven et al., 2013; Van Boven & Loewenstein, 2005). Because people see the helper's emotions as especially informative about the moral status of a prosocial act (Barasch et al., 2014; Montealegre et al., 2021), they may use their own anticipated emotional states to infer how the helper was feeling—and thus, how much moral credit he or she deserves.

Distress Versus Empathic Concern

What morally relevant states might observers be inferring? Two affective states central to the literature on prosocial behavior are distress and empathic concern (Batson et al., 1983). Distress is defined as a feeling of "personal uneasiness and discomfort" engendered by the victim's need or suffering, whereas empathic concern is defined as a feeling of "sensitivity and concern" directed toward the victim (Batson et al., 1983, Batson, 1987). Empathic concern is often measured by the endorsement of more emotionally positive items (warm, tender, sympathetic, softhearted, moved, compassionate), and distress by the endorsement of more emotionally negative items (worried, upset, troubled, perturbed, grieved, disturbed, alarmed, distressed).

When it comes to actual helping behavior, a potential helper's feeling of distress versus empathic concern is linked to different outcomes; people feeling distressed tend to help only if no other means of reducing their negative emotion is available (e.g., distancing themselves from the situation). Those who feel empathic concern, on the other hand, are more likely to help even if they have the option to distance themselves from the victim or improve their mood in other ways (see Batson, 2010; Batson et al., 2002, for reviews). When it comes to predicting one's own emotional states or judging those of others, however, the distinction between distress and empathic concern is less clear. Shaw et al. (1994) found that when forecasting their own emotions, participants did not clearly differentiate distress and empathic concern, and Barasch et al. (2014) found that when judging others who helped, perceived distress and empathic concern both led to more positive moral character judgments. To capture as broad a range of affect as possible (and, in particular, to capture both positive and negative affect), we measured both empathic concern and distress in the present studies. Based on the prior research, we expected that both inferred distress and concern would be associated with more positive moral judgments, though we were agnostic about whether the associations would be stronger for distress or empathic concern.

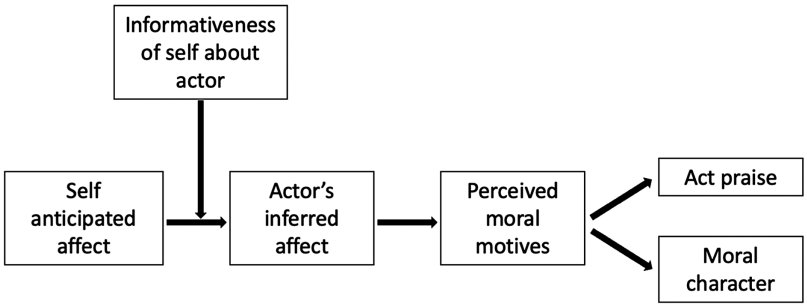
The Current Research

Our overall hypothesis is that observers' anticipated affect—that is, the emotion that they expect to feel in a situation where helping is possible—will predict moral praise for the act of helping as well as moral character judgments of the helpers. Specifically, we expect that actors who help in situations where observers expect to feel more emotion will be praised more and will be seen as having more positive moral character. We hypothesize that this is because observers use their own anticipated affect to infer the affect experienced by the actor, and that greater perceptions of actor affect will lead to actors being seen as more morally motivated. Perceived moral motivation, in turn, is hypothesized to lead to more praise for helping and more positive perceptions of the actor's moral character (see Figure 1).

There are two general ways of testing our hypotheses: a between-participants approach and a within-participants approach. In the more common within-participants approach, every participant makes judgments for every variable in the causal chain. This has the advantage of establishing that individual-level variation in (for example) anticipated distress is associated with individual-level variation in praise judgments. It does, however, have a disadvantage. By asking the same participant to respond to every item, it may make salient relationships between questions that would not otherwise occur to participants, thereby affecting their responses. This disadvantage is avoided in a between-participants approach, in which one group of participants rates a set of scenarios on

Figure 1

Conceptual Diagram Showing the Hypothesized Causal Path From Anticipated Affect to Inferred Actor Affect, Moral Motives, and Praise/Moral Character Judgments



Note. The link between anticipated and inferred affect is moderated by whether observers think their anticipated affect can be used to infer the actor's.

anticipated affect, and a separate group rates them on the dependent variables (moral motives, praise, and character judgments). The statistical test of the hypothesis in this case is whether scenario-level variation in anticipated affect is associated with scenario-level variation in the dependent variables. Of course, this approach has the drawback that it does not test these relationships at the individual level. Employing both approaches then enabled us to mitigate the limitations characteristic of each approach. Studies 1 and 2 used a between-participants approach, whereas Studies 3 and 4 used a within-participants approach.

Between-Participants Studies (Studies 1 and 2)

We first constructed a list of 20 situations in which helping was possible (see the Scenario Development section). We then measured the anticipated affect for each scenario by asking participants how distressed and empathically concerned they would feel if they were confronted with each scenario. Next, to be able to control for the perceived positive consequences of helping, we paired each scenario with a helping outcome and asked participants to rate how beneficial helping would be in each case. We then proceeded to the studies that tested our main hypotheses. In Study 1, we asked participants to evaluate the moral character and praiseworthiness of people who helped in each scenario and tested whether helpers in scenarios that had previously been rated as more affectively arousing were evaluated more positively. In Study 2, we asked participants to evaluate the moral character and praiseworthiness of the helpers, as well as their perceived motives. We then tested whether helpers in more affectively arousing scenarios were seen more positively because of their perceived moral motives. We expected this to be the case even when statistically controlling for the perceived positive consequences of the helping act.

Within-Participants Studies (Studies 3 and 4)

Study 3 used the same 20 helping situations that we developed for Studies 1 and 2. Participants were randomly assigned to rate one of the 20 situations on each component of our hypothesized causal chain: (self) anticipated affect, helper affect, helper motives, and action praise/actor character judgments. In this study, the statistical test of the hypothesis is whether the overall path coefficient differs from zero. As in Studies 1 and 2, we also controlled for the perceived positive consequences of helping (as rated by each participant).

Study 4 used the same design as Study 3, with one addition: Participants were randomly assigned to a control condition in which they saw the same description of helping as in Study 3, or to a condition in which they read that due to a brain condition, the helper did not feel "strong emotional impulses that signal what is morally right or wrong." We predicted that in this condition, participants would be less likely to use their own anticipated affect to infer the target's emotionality, and that this would have downstream consequences for impressions of the target's moral motives, moral character, and praiseworthiness. The statistical test of this hypothesis is whether the paths corresponding to the hypothesized causal chain differ between conditions in a multigroup structural equation model.

Open Practices Statement

Data from all studies, R code to reproduce all analyses, preregistrations, and Supplemental Materials are available at https://researchbox.org/429.

Scenario Development

We first needed to construct a set of scenarios in which helping was possible. Our criteria were: that scenarios needed to be mundane enough that participants could reasonably imagine themselves encountering them; that the need for help was salient; that they varied in how much affect they evoked; and that we could realistically modify them to include helping by a third party. Rather than inventing our own scenarios, we (as much as possible) adapted scenarios used in prior research. Four of our scenarios were adapted from Yudkin et al.'s (2019) helping actions. We also adapted three scenarios from Kahneman and Ritov's (1994) "headlines" study and three scenarios from McManus et al.'s (2020) moral obligation study. We created one other scenario by combining elements from both Yudkin et al. (2019) and McManus et al. (2020) scenarios regarding paying for another person at a cafe. Last, nine more scenarios were independently constructed by us after reading

through these three sources (Kahneman & Ritov, 1994; McManus et al., 2020; Yudkin et al., 2019) for general trends. This approach yielded 20 total scenarios (for the full text of each scenario, including source adaptation, see Table 1).

The 20 scenarios were intentionally varied across dimensions highlighted in the moral perception literature. We varied the identity of the target in need across dimensions of gender and age (e.g., "elderly man," "small girl"), human versus animal species (e.g., "lost dog," "several Australian mammal species"), and number (e.g., "someone struggling," "group of teenagers"). We also varied the nature of the scenarios with regard to immediacy (e.g., "a new hire at work," "watch[ing] the news"), urgency (e.g., "an ad asking to donate to Wikipedia," "somebody on the sidewalk nearby in distress"), and whether or not there was a direct request for assistance (e.g., "someone … with a sign that says 'I'm hungry'," "someone struggling to carry a heavy box"). In order to keep scenarios short and realistic, only a subset of these dimensions were made salient in every scenario, that is, only the intuitively relevant details were included. For example, the age of the target in need is only mentioned in the scenario where the helper gives their seat to an elderly man on the subway or sees an elderly woman trying to cross the street.

Scenario Ratings: Anticipated Affect

For Studies 1 and 2, we needed to determine how much affect each scenario evoked. We therefore asked participants to rate how much distress and empathic concern they would feel if confronted with each situation. We aimed to have about 50 participants rate each scenario. As each participant rated only five of the 20 scenarios, we recruited 200 participants in total.

Method

Participants

We recruited 200 U.S. adult participants from Amazon's Mechanical Turk using CloudResearch and enabled the CloudResearch quality settings "Verify Worker Country Location," "Suspicious Geocode Block," and "Duplicate IP Block." Of these, 173 passed our attention check (see below) and were included in the sample (M_{age} = 35.9, range = 18–66; 65.3% male, 34.1% female, 0.6% undisclosed). Participants were paid $.45 for their time.

Procedure

Participants were randomly assigned to rate one of four subsets of five scenarios of the 20 total shown in Table 1. After reading each scenario, participants were asked to indicate how much they would feel each of the 14 emotions using 7-point Likert scales (1 = *not at all*, 7 = *very much*). The 14 emotion items were taken from Batson et al.'s (1983) scales measuring empathic concern (*warm, tender, sympathetic, softhearted, moved, compassionate*) and distress (*worried,*

Table 1
Full Text of Each Scenario

Scenario	Full text
Article[a]	Imagine that you read an article "Several Australian mammal species nearly wiped out by hunters."
Car malfunction	Imagine that you are driving and see a smoking car on the side of the road. It is apparent that the car's engine had begun to malfunction.
Cashier[b,c]	Imagine that you are waiting in line in a cafe, and the person in front of you is ordering at the cashier when they realize that they have left their wallet at home.
Documentary	Imagine that you have just watched a moving documentary about poverty in inner-city America.
Dropped cash	Imagine that you see someone accidentally drop a $20 bill without noticing.
Elderly woman[b]	Imagine that you see an elderly woman struggling to cross the street.
Facebook confession	Imagine that you are scrolling through Facebook during a work break. Right before you close the app, you see a post from someone you don't know well who goes to your school with you. The post is a detailed confessional about how they are having a hard time fitting in and making friends.
Go fund me[c]	Imagine that you are scrolling through Facebook during a work break. Right before you closed the app, you see two similar posts that were shared by a friend in which people were using "Go Fund Me" to ask for donations to help pay for flood damage to their homes after a severe thunderstorm.
Heavy box[c]	Imagine that you see someone struggling to carry a heavy box.
Hungry sign[c]	Imagine that you see someone sitting on the sidewalk with a sign that says "I'm hungry."
Ice skates	Imagine that you are at an ice rink, and you see a small girl having trouble putting on her ice skates.
Kittens	Imagine that you find a litter of abandoned kittens in a dumpster.
Lost dog[b]	Imagine that you see a lost dog with a collar on.
New employee	Imagine that a new hire at work is having a hard time figuring out how to find her office.
Newspaper[a]	Imagine that you read a statistic in a newspaper article describing the number of deaths due to leukemia: "Approximately every 9 minutes, someone in the United States dies from a blood cancer."
Red cross	Imagine that a group of teenagers are fundraising by soliciting for donations for the Red Cross outside of the building you are about to walk into.
Subway man[b]	Imagine that you are riding the subway to work and are seated comfortably. You see an elderly man board the train and look for a seat, but all the seats are taken.
Waving cars[b]	Imagine that you are driving home from work and see somebody on the sidewalk nearby in distress and waving down cars.
Wikipedia	Imagine that you log onto Wikipedia and see an ad asking to donate to Wikipedia.
Wildfire[a]	Imagine that you watch the news and see the aftermath of a series of devastating wildfires.

[a] Scenario item(s) adapted from Kahneman and Ritov (1994). [b] Scenario item(s) adapted from Yudkin et al. (2019). [c] Scenario item(s) adapted from McManus et al. (2020).

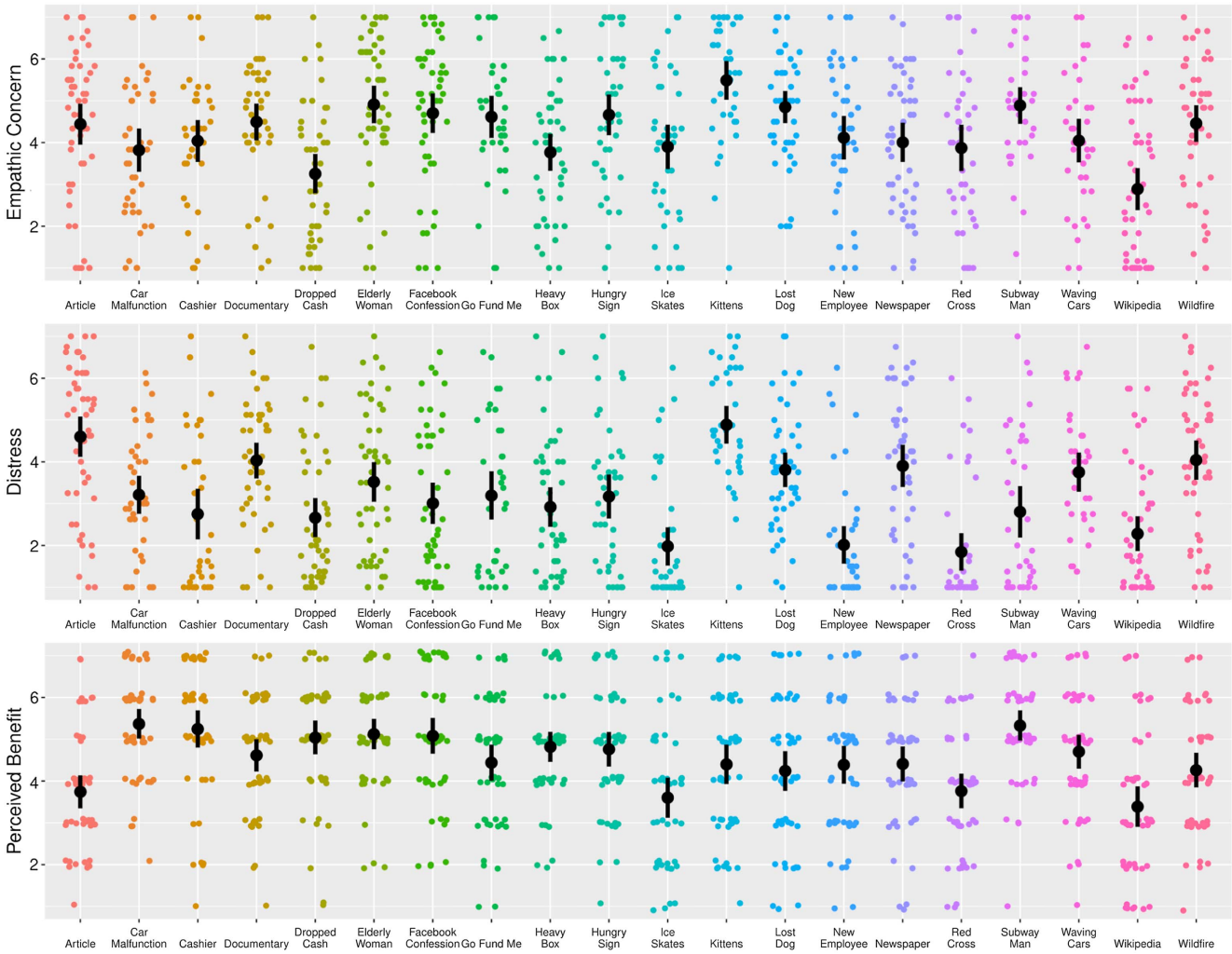
upset, troubled, perturbed, grieved, disturbed, alarmed, distressed). These emotion items were displayed in a random order for each participant. After the five scenarios, participants were shown an attention check that was formatted similarly to the scenarios but read, "This is an attention check. Please choose '1' for every question below." Participants were coded as passing the attention check if they chose "1" for every emotion item in that block. Finally, participants completed demographic questions (age, gender, and race).

Results

We first confirmed that the structure of Batson et al.'s (1983) 14-item empathic concern and distress inventory was as expected by running a principal components analysis with varimax rotation on the scenario-level aggregated item scores. The first two components explained 93% of the total variance, and each item loaded highly onto its expected component: empathic concern (highest loading: compassionate = 0.93; lowest loading: moved = 0.76); distress (highest loading: disturbed = 0.96; lowest loading: perturbed = 0.89). Thus, participants differentiated between empathic concern and distress in their anticipated affect ratings. We therefore computed composite empathic concern and distress scores for each scenario by averaging ratings for the distress and empathic concern items for each scenario across participants (all αs ≥ .89). Distress and empathic concern scores for each scenario are shown in Figure 2 (top two panels); per-scenario means and standard deviations are available in

Figure 2

Empathic Concern, Distress, and Benefit Ratings by Scenario



Note. Colored dots indicate individual participant ratings for each scenario. Black dots indicate the mean across all participants' ratings for a scenario, with black line indicating the 95% confidence interval. See the online article for the color version of this figure.

Supplemental Table S2. There was more agreement between raters on ratings of distress (ICC[3] = .38) than empathic concern (ICC[3] = .23).[1] Consistent with past research (e.g., Batson et al., 1983), scenario-level distress and empathic concern were positively correlated, $r(20) = .58$, $p = .01$.

Scenario Ratings: Perceived Benefits

Although outcomes seem to matter less for praise than blame judgments (Anderson et al., 2020), they do still matter. It is, therefore, useful to know the perceived benefits of helping in each scenario so that we can test whether any effects of anticipated affect on moral judgments persist even when controlling for the positive consequences of the act. Therefore, we asked a separate group of participants to rate our scenarios on how beneficial helping would be. We again aimed to have 50 participants rate each scenario, but because in this study we asked participants to rate 10 scenarios rather than five, we recruited 100 participants total.

Method

Participants

We aimed to recruit 100 U.S. adult participants from Prolific.co, a prescreened online labor market for behavioral research. We recruited a total of 103 participants, of whom 101 passed our attention check (which was identical to that in the previous studies) and were included in the sample ($M_{age} = 34.8$, range = 18–77; 49% male, 51% female). Participants were paid £.35 (about $.50 U.S.) for their time.

Procedure

We modified each scenario to include helping behavior (e.g., "A new hire at work is having a hard time figuring out how to find her office. Noticing this, [actor] offers to help show her where it is."; see Supplemental Table S3, for full text and Supplemental Appendix A, for the full list of actor names). Participants were randomly assigned to see one of two subsets of 10 scenarios.

Our measure of perceived benefits is based on one developed by Yudkin et al. (2019). After each scenario, participants were asked, "In your opinion, how much of a positive effect would the act described above have overall?" on 1–7 Likert scales (1 = *the LEAST positive effect overall*; 7 = *the MOST positive effect overall*). To avoid ceiling effects, participants were shown a reminder before each rating that "while every good deed may have some positive effects, some might have greater positive effects than others."

Results

We averaged the benefit ratings for each scenario across participants. Mean perceived benefits ratings ranged from 3.31 ($SD = 1.67$) for the Wikipedia scenario to 5.33 ($SD = 1.26$) for the Car Malfunction scenario. Benefit ratings for each scenario are shown in Figure 2 (bottom panel) and Supplemental Table S3.

Discussion

We use these scenario-level benefit ratings to control for perceived positive consequences of helping in Studies 1 and 2. This is important because the analyses in these studies are correlational at the scenario level. Perceived positive consequences of helping could plausibly correlate both with affective responses to a situation as well as moral praise and character judgments of helpers. Controlling for perceived benefits ensures that any relationship between anticipated affect and moral judgments is not due to this potential confound.

Study 1

In Study 1, we tested whether helpers in more affectively arousing situations received more praise and more positive character judgments than helpers in less affectively arousing situations. We hypothesized that compared to helpers in less affectively arousing situations, people who helped in more affectively arousing situations would be praised more and would be seen as having better moral character.

We preregistered our planned sample size, predictions, and analyses at https://aspredicted.org/blind.php?x=ug3ww7. All analyses reported here can therefore be considered confirmatory unless otherwise specified.

Method

Participants

We aimed to recruit 300 U.S. adult participants from Prolific.co, a prescreened online labor market for behavioral research. We recruited a total of 301 participants, of whom 286 passed our preregistered attention check (which was identical to the one used in the two previous rounds of data collection). A further two participants indicated that they were under 18. Although we had not preregistered an age exclusion, our ethics approval did not allow us to use data from minors, and therefore, we excluded their data as well. The final sample thus consisted of 284 participants ($M_{age} = 34.3$, range = 18–75; 42.8% male, 55.1% female, 2.1% other). Participants were paid £.35 (about $.50 U.S.) for their time.

Procedure

We used the same scenarios as in the "Perceived Benefits" study, such that every scenario described a protagonist who provided help. Participants were shown 10 of the 20 possible scenarios (randomly selected). After reading each scenario, participants were asked to answer four questions (all on 7-point scales). One question asked them to evaluate the helping act: "How morally praiseworthy would it be to perform the act described above?" (1 = *the LEAST praiseworthy a good deed could be*, 7 = *the MOST praiseworthy a good deed could be*). The remaining three asked them to evaluate the actor: "Based on what you read, how moral is [actor]?" (1 = *not at*

[1] Both values are below the standard cutoff for "low" agreement of .5 (Liljequist et al., 2019). Low agreement between participants is not a theoretical problem. Our account is that observers make moral judgments based in part on their anticipated affective responses, not that anticipated affective responses will be the same across individuals. Indeed, Studies 3 and 4, which are within participants, assume that people will vary in their anticipated affect. In this way, the data here are unlike those for which interrater agreement is usually computed, where the ideal is perfect agreement between raters. Furthermore, in Studies 1 and 2, we use the scenario-level aggregates, which "average out" rater effects. However, the difference in agreement between raters for distress versus empathic concern is noteworthy, and we return to this in the Discussion section of Study 2.

all moral, 7 = *very moral*); "Based on what you read, do you think [actor] is mainly a good person or a bad person?" (1 = *mainly a bad person*, 7 = *mainly a good person*); "Based on what you read, do you think [actor] has good moral standards?" (1 = *not at all*, 7 = *completely*). Question order was randomized per participant. Finally, participants completed the attention check questions and demographics.

Results

Our primary hypotheses were that scenario-level anticipated distress and empathic concern would predict greater praise for helping and more positive evaluations of helpers. We used the previously collected ratings of anticipated distress, empathic concern, and perceived benefits as the scenario-level variables. Average praise and character ratings for each scenario are shown in Supplemental Table S4.

The three actor evaluation items from the present study were combined into a composite measure of perceived moral character ($\alpha = .93$). We then followed our preregistered analysis plan and fit four separate mixed-effects models that modeled praise and perceived moral character as a function of scenario-level anticipated distress, empathic concern, and perceived benefits (these were treated as fixed effects). We also included randomly varying intercepts for scenarios and participants to account for nonindependence of ratings of the same scenario and from the same participant (note that these effects were crossed, as each participant only rated half of the total scenarios). These analyses revealed that as predicted, the more scenarios had previously been rated as evoking distress and empathic concern, the more helping was praised, and the more helpers were seen as having good moral character (see Table 2).

Exploratory Analyses (Not Preregistered)

We also tested whether the results were robust to removing the control for perceived benefits. Removing perceived benefits from the models reported above did not change the direction or significance of any of the results. These results are reported in detail in the Supplemental Table S6.

Discussion

These findings support our initial hypothesis that people who helped in more affectively arousing situations would be praised more and would be seen as having better moral character. Although these findings show that helpers are evaluated more positively when they help in affectively arousing situations, they do not explain what inferences are driving these moral judgments. We turn to this question in Study 2.

Study 2

Having found that helpers are evaluated more positively when they help in affectively arousing situations, we next turned to the second part of our account, which is that this is because of inferences about helpers' moral motives. We hypothesized that helpers in more affectively arousing situations would be seen as more morally motivated, which would statistically mediate the link between anticipated affect and praise and character judgments. We preregistered our planned sample size, predictions, and analyses at https://aspredicted.org/8WS_KKL. All analyses reported here can therefore be considered confirmatory unless otherwise specified.

Method

Participants

Each participant rated one scenario in Study 2. In a previous version of this study (which is reported in the Supplemental Material), we recruited 1,000 participants and found inconsistent support for our hypotheses, though all effects were in the expected direction. With $N = 1{,}000$, each scenario was rated by 50 participants on average (as opposed to 150 on average in Study 1), so we thought the mixed results might be explained by lower statistical power due to the lower per-scenario sample size. In this version of the study, we therefore requested 3,000 U.S. adult participants from Prolific.co to match the per-scenario sample size of Study 1. We recruited 3,017 participants, of whom 2,867 passed our preregistered attention check (which was identical to the one used in Study 1). A further four participants indicated that they were under 18. As preregistered, we excluded these participants as well. The final sample thus included 2,863 participants (M_{age} = 40.65, range = 18–94; 48.7% male, 48.4% female, 2.5% nonconforming/not listed, 0.4% nondisclosed). Participants were paid £.15 (about $.20 U.S.) for their time.

Procedure

We used the same scenarios as in Study 1, such that every scenario described a protagonist who provided help. Because in this study participants completed more items for each scenario, each participant was randomly assigned to see only one of the 20 possible scenarios. After reading the scenario, participants completed the same four act and actor evaluation items from Study 1. They were also asked to indicate how much the actor's behavior was based on "genuine moral concern," "moral principle," "a genuine moral stand," "personal self-interest," "what was good for [actor] personally," "selfish reasons," "what was reasonable," "what was rational," and "a pragmatic stand" (all on 1–7 scales anchored by *definitely no* and *definitely yes*). These

Table 2
Results of Mixed-Effects Models Predicting Praise for Helpful Acts and Evaluations of Helper Character From Scenario-Level Anticipated Distress and Empathic Concern (Separately)

	Distress				Empathic concern			
Outcome	*b* (*SE*)	*t*	*df*	*p*	*b* (*SE*)	*t*	*df*	*p*
Act praise	.16 (.05)	3.40	16.99	.003	.19 (.08)	2.50	16.87	.02
Character evaluation	.12 (.04)	3.22	16.99	.004	.14 (.06)	2.43	16.93	.03

Note. All models control for scenario-level perceived benefits of helping. *SE* = standard error.

items were developed by Uhlmann et al. (2013). The first six measured perceived moral motives, and the final three measured perceived pragmatic motives (note that half the moral motives items are reverse-scored). Analysis of data from the previous version of this study suggested that three of the six moral motives items performed poorly, as evidenced by poor item characteristic curves (see Supplemental Figure S1). Thus, in our preregistration for Study 2, we specified that we would measure moral motives using only the three items that performed well in the previous version: how much the actor's behavior was based on "genuine moral concern," "moral principle," and "a genuine moral stand." As a robustness check, we report the results of our analyses using the full six-item scale in the Supplemental Material (see Supplemental Table S9).

Question order was randomized per participant with the constraint that the moral motives items always appeared together as a block. Finally, participants completed the attention check questions and demographics.

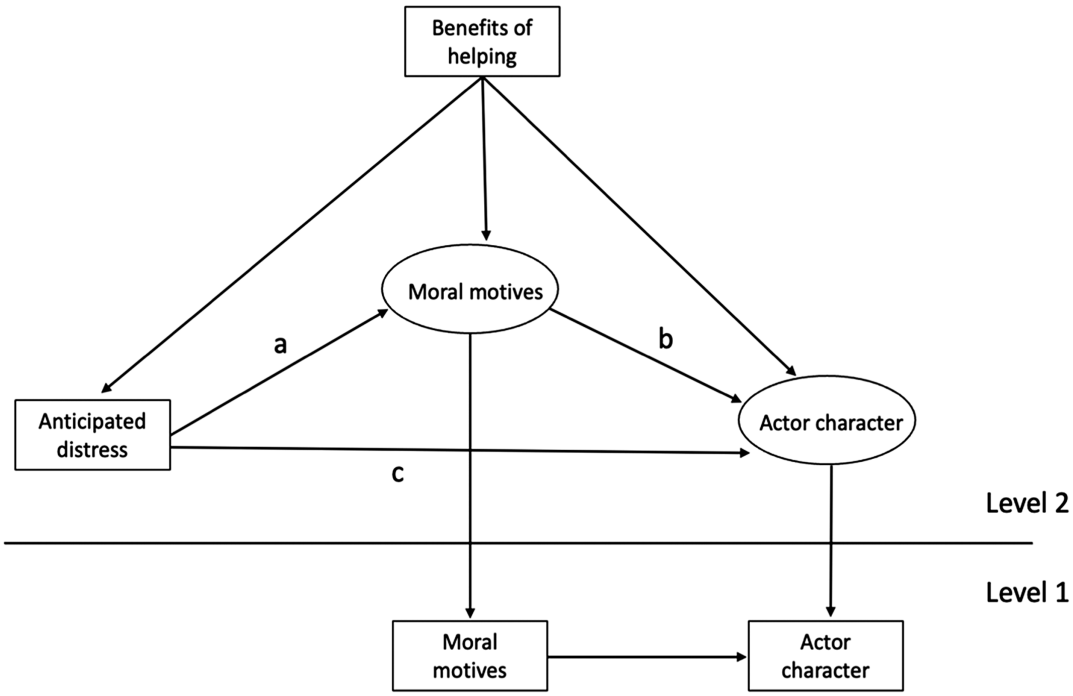
Results

As in Study 1, we averaged the three actor evaluation items to produce a composite measure of perceived moral character (α = .91). We also combined the three moral motives items into a moral motives composite (α = .88), and the three pragmatic motives items into a pragmatic motives composite (α = .68). Average moral motives, pragmatic motives, praise, and character ratings for each scenario are shown in Supplemental Table S7.

Our hypothesized causal path was that scenario-level anticipated emotion would lead to greater perceived moral motivation of the actor, which in turn would lead to greater praise and more positive evaluations of the actor's character. Statistically, this is operationalized as a significant indirect path from scenario-level anticipated distress/empathic concern to perceived moral motives to praise/character judgments. Our preregistered analysis plan was to test these indirect paths using multilevel structural equation modeling (SEM; Preacher et al., 2010) to account for the fact that participants were nested within scenarios (i.e., multiple participants rated the same scenario), and that the independent variable (anticipated affect) was measured at the scenario level, whereas the mediator and dependent measure were measured at the individual level (i.e., a 2-1-1 multilevel mediation model; see Krull & MacKinnon, 2001).

We followed the recommendations in Preacher et al. (2010) for a 2-1-1 design and fit structural equation models that simultaneously modeled Level 1 (i.e., participant level) paths between the mediator (moral motives) and dependent variable (praise/character judgments) and Level 2 (i.e., scenario level) paths from the independent variable (empathic concern/distress) to the mediator to the dependent variables, as well as direct effects of the independent on the dependent variables (these correspond to what are often called the *a*, *b*, and *c* paths in mediation models). All models also included Level 2 paths between perceived benefits of helping and each variable in the causal chain (i.e., anticipated affect, moral motives, and praise/character judgments) in order to control for possible confounding effects of perceived benefits (see MacKinnon & Pirlott, 2015). Indirect effects were only measured at Level 2 (because the independent variable only varied at the scenario level). An example model is shown graphically in Figure 3.

Figure 3
One of Four Multilevel Mediation Models



Note. Moral motives and character/praise are measured at the individual level (Level 1). Anticipated concern/distress is measured at the scenario level (Level 2). Indirect effects are estimated based on paths (a, b, and c paths) at Level 2. All models include effects of perceived benefits of helping at Level 2.

Table 3
Estimated Indirect Effects of Scenario-Level Anticipated Distress and Empathic Concern on Praise and Character Judgments via Perceived Moral Motives

	Distress			Empathic concern		
Outcome	Estimate (*SE*)	*Z*	*p*	Estimate (*SE*)	*Z*	*p*
Act praise	.19 (.07)	2.94	.003	.17 (.09)	2.01	.05
Character evaluation	.17 (.05)	3.59	<.001	.12 (.06)	2.03	.04

Note. All models control for scenario-level perceived benefits of helping. *SE* = standard error.

Tests of the indirect effects were consistent with our predictions. As Table 3 shows, we observed significant indirect effects of anticipated distress and empathic concern on both actor and character judgments.

We also fit the same models with pragmatic motives, rather than moral motives, as the mediator. As expected, there were no significant indirect effects in any of the four models (all *p*s > .40).

Secondary Analyses

As a conceptual replication of Study 1, we tested the total effects of anticipated distress and concern on praise and character judgments using the same multilevel mediation models. Because the total effect consists of the (Level 2) direct + indirect effects, it is conceptually equivalent to the models in Study 1 that tested the effects of scenario-level anticipated distress and concern on act and character judgments. For anticipated distress, the total effect was significant for praise ($Z = 3.84$, $p < .001$) and character judgments ($Z = 2.96$, $p = .003$). For anticipated empathic concern, the total effect was significant for character judgments ($Z = 2.03$, $p = .04$) but not for praise ($Z = 1.86$, $p = .06$).

We also tested whether the results were robust to omitting the control for perceived benefits. After removing perceived benefits from the models reported above, the total effect of empathic concern on praise became significant ($Z = 2.25$, $p = .02$). There were no other changes in the direction or significance of any results. These results are reported in detail in Supplemental Table S10.

Discussion

This study tested if the effects of anticipated distress and empathic concern on praise and character judgments resulted from inferences about the actor's moral motives. We predicted that actors who helped in more affectively arousing situations would be seen more positively in part because of inferences about their motives—specifically, because helpers in affectively arousing situations would be seen as more morally motivated. Consistent with our predictions, we found significant indirect effects of anticipated affect on both actor and character judgments via moral motives. Taken together, then, our findings suggest that helpers in more affectively arousing situations may be seen more positively in part because they are seen as more morally motivated.

The evidence for this is stronger for distress than empathic concern; for the latter, the key tests of indirect effects were only just statistically significant at the .05 level. The weaker effects for empathic concern may result from the fact that there is lower agreement between participants on how much empathic concern they would feel in each scenario (ICC[3] = .23) than for distress (ICC[3] = .38). Low agreement between participants is not a theoretical problem: our account is that observers make moral judgments based in part on their anticipated affective responses, not that anticipated affective responses will be the same across individuals. But when interrater agreement is low, scenario-level estimates will be noisy, and statistical power will be lowered. This lower power is a weakness of the between-participants design, where the indirect effects are estimated at the scenario level. In Study 3, we, therefore, move to a within-participants design that maximizes power by measuring all variables for every participant.

Study 3

In Study 3, we moved from the between-participants design of Studies 1 and 2 to a design in which participants rated a single randomly selected scenario on anticipated affect, helper moral and pragmatic motives, and praise and character judgments. We also tested an additional part of the hypothesized causal chain that was not tested in Studies 1 and 2: perceived helper affect. Our account holds that perceivers base their perceptions of a helper's emotional state on their own anticipated emotional reactions to the situation, and that these emotion perceptions then have downstream consequences for perceived moral motives and, ultimately, praise and character judgments. We, therefore, asked participants to indicate what emotions they thought the helper was experiencing in the situation (in addition to predicting how they themselves would feel). As in Studies 1 and 2, we also controlled for perceived benefits of helping.[2] This design allowed us to test the entire hypothesized causal process in a single model, which we did using SEM. An example model is shown in Figure 4.

We preregistered our planned sample size, predictions, and analyses at https://aspredicted.org/TJB_91N. All analyses reported here can therefore be considered confirmatory, unless otherwise specified.

Method

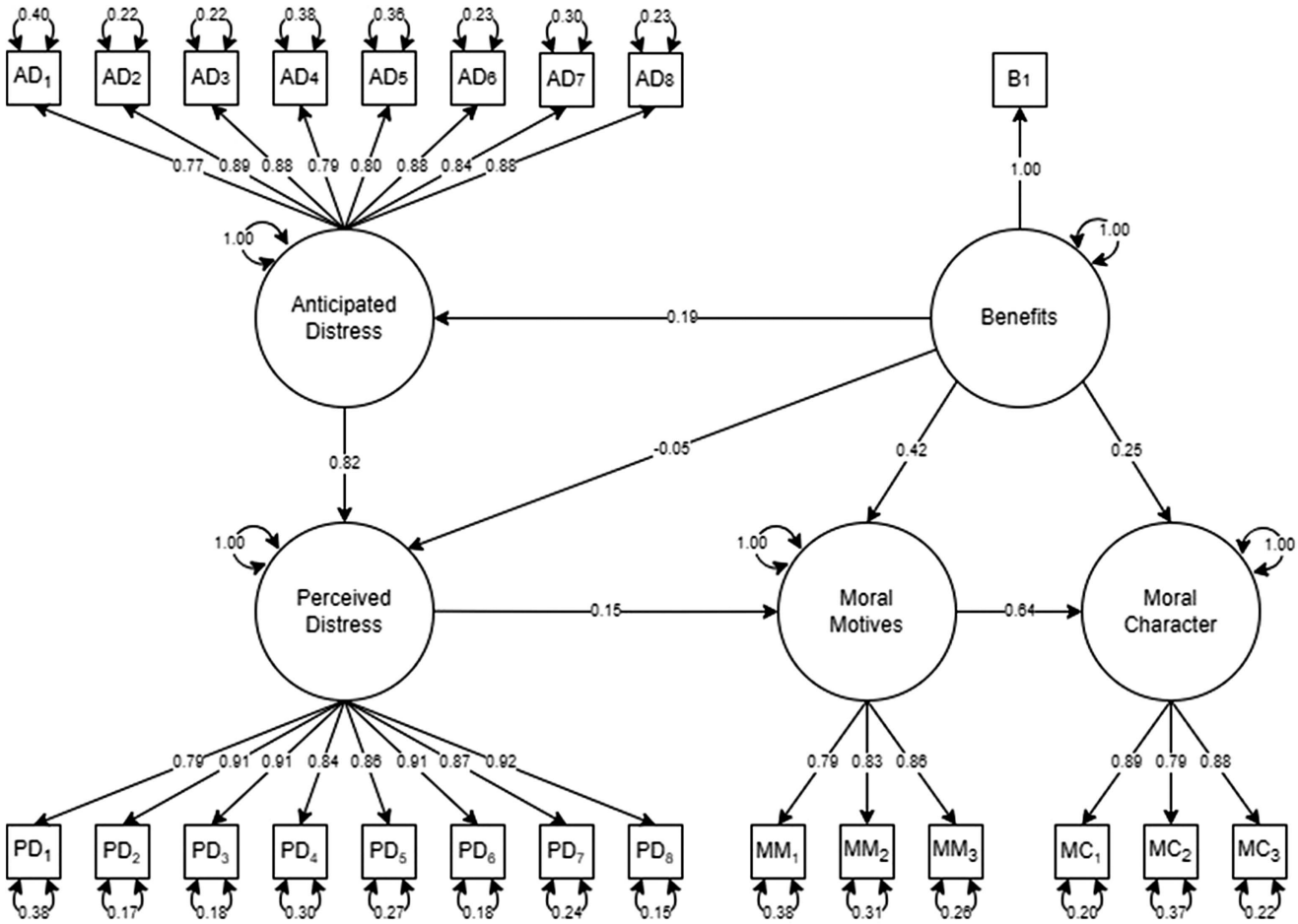
Participants

We recruited a total of 1,000 U.S. adult participants from Prolific.co, a prescreened online labor market for behavioral research. Of these, 962 passed our preregistered attention check (participants were asked to select a grammatically correct word that

[2] In this study, we preregistered models that allowed the benefits of helping to affect both perceived moral motives and praise/character judgments. A reviewer subsequently suggested that we allow perceived benefits to affect anticipated and perceived affect as well. The models reported below include all these paths, even though only the first two were preregistered. Models omitting perceived benefits are reported in the Supplemental Material.

Figure 4

Structural Equation Model for Anticipated Distress and Moral Character Judgments



Note. AD = anticipated distress; PD = perceived distress; MM = moral motives; MC = moral character.

completes the sentence: "What time ____ bus leave?"). A further three participants either indicated that they were under 18 or failed to disclose their age. We had not preregistered an age exclusion[3] but our ethics approval did not allow us to use data from minors. We therefore excluded their data as well (though these age reports were likely keystroke errors, as it would be unlikely for individuals under the reported age of two to be able to take an online survey). The final sample thus consisted of 959 participants (M_{age} = 37.7, range = 18–79; 48.5% male, 48.9% female, 2.6% other). Participants were paid £.50 (about $.65 U.S.) for their time.

Procedure

Participants first read one randomly chosen scenario of the 20 shown in Table 1. Each scenario asked the participant to imagine themselves in a situation where helping was possible (but did not describe a helping behavior).

Anticipated Affect. After reading the scenario, participants were asked to indicate how much they would feel each of the 14 emotions using 7-point Likert scales (1 = *not at all*, 7 = *very much*). (This was the same procedure as in the "Scenario Ratings: Vicarious Affect" study). The 14 emotion items were taken from Batson et al.'s (1983) scales measuring empathic concern (warm, tender, sympathetic, softhearted, moved, compassionate) and distress (worried, upset, troubled, perturbed, grieved, disturbed, alarmed, distressed).

Next, participants read the same scenario but with the addition of a protagonist who helped (these are the same scenarios as were used in Studies 1 and 2; the full text of each is shown in Supplemental Table S3). After reading the scenario, participants were asked to respond (using 7-point Likert scales) to the following measures in random order.

Perceived Helper Emotions. Participants were asked to indicate how much they thought the protagonist felt each of the 14 emotions from Batson et al.'s (1983) scales when deciding to help. These items were presented as a block but randomized within the block.

Moral and Pragmatic Motives. These were the same motives questions used in Study 2 and first developed by Uhlmann et al. (2013). Six measured perceived moral motives, and three measured perceived pragmatic motives. These items were presented as a block but

[3] Study 3 was run before Study 2, in which we preregistered an age exclusion.

Table 4
Measures Used in Study 3

Measures	*M* (*SD*)	α	ICC
Perceived empathic concern	4.28 (1.56)	.92	0.15
Anticipated empathic concern	5.48 (1.13)	.87	0.11
Anticipated distress	3.17 (1.70)	.95	0.43
Perceived distress	3.24 (1.83)	.96	0.53
Moral motives	5.67 (1.09)	.86	0.10
Moral character	5.94 (0.89)	.89	0.06
Perceived benefits	5.05 (1.20)		0.09
Moral praise	4.85 (1.21)		0.10

Note. ICC = intraclass correlation coefficient.

randomized within the block. In our preregistration, we specified that we would measure moral motives using only the three items that performed well in previous studies. As a robustness check, we report the results of our analyses using the full six-item scale in the Supplemental Material (see Supplemental Tables S8 and S9). All results were consistent with the findings reported below.

Perceived Benefits. This was a single-item measure of perceived benefits of helping. It was identical to that used in the "Scenario Ratings: Perceived Benefits" study.

Act and Actor Judgments. These were the same four questions used in Studies 1 and 2; one asked participants to evaluate the praiseworthiness of the act, and the remaining three asked about the moral character of the actor (i.e., the helper in the scenario).

Last, participants completed the attention check question and demographics.

Results

We report descriptive statistics and intraclass correlation coefficients (ICCs) in Table 4. We performed SEM in R (Version 3.6.0; R Core Team, 2019) using the lavaan package (Rosseel, 2012). Following the best practice recommendations (Morrison et al., 2017), we tested a measurement model with all latent constructs prior to testing structural models. Fit statistics indicated that this measurement model (comparative fit index [CFI] = 0.90, standardized root-mean-square residual [SRMR] = 0.07, root·mean-square error of approximation [RMSEA] = 0.08, χ = 3576.40, *df* = 540) passed liberal thresholds for model fit (CFI > 0.90 from Schumacker & Lomax, 2010; SRMR < .08 from Hu & Bentler, 1999; RMSEA < .08 as indicating medicore fit, see MacCallum et al., 1996). Overall, these fit statistics suggested that the measurement model was not an excellent fit for the data but was sufficient to allow us to continue with the test of our hypothesized pathway.

After estimating the measurement model, we fit models that included our predicted pathways between the constructs. For distress, we tested two separate models: one for moral character and one for moral praise. Following this, we tested the effect of empathic concern, again using separate models for moral character and moral praise.

As specified in our preregistration, we intended to use cluster-robust standard errors to account for between-scenario variance in ratings. However, models with cluster-robust standard errors failed to converge. Thus, we present all SEM models without these standard errors. As an alternate analysis that accounts for between-scenario variance, we also tested our main hypothesized pathway using serial mediation models with cluster-robust standard errors. The results of these analyses (reported in full in the Supplemental Materials) replicate the findings reported below.

Summary of the Models

For all four models, the hypothesized indirect paths (from anticipated emotion to perceived emotion, to moral motives, and finally to praise/character judgments) were statistically significant. Fit statistics and indirect path tests for each model are shown in Table 5. Though indirect path tests were statistically significant for every model, indirect effects were three to four times larger for moral character compared to moral praise. Figure 4 shows the full diagram for one of the four models (distress/moral character); diagrams for the remaining three are available in the Supplemental Materials (see Supplemental Figures S2, S3, and S4). All paths in the models were preregistered, except for the paths allowing the perceived benefit to predict perceived distress and anticipated distress (these were added at the request of an anonymous reviewer and did not impact the statistical significance of our main pathways).

Alternative Models

As a test of alternative causal models, we fit several additional models, each with an additional pathway. The fit of each of these models was then compared to the corresponding original model (see Table 6). Some of these models were a better fit for the data than our

Table 5
Indirect Paths From Structural Equation Models in Study 3

	Predicted pathway			Model fit				
Outcome variable	*B* (*SE*)	*Z*	*p*	RMSEA	SRMR	CFI	χ^2	*df*
Distress								
Act praise	0.04 (0.01)	4.14	<.001	0.09	0.05	0.92	1693.59	184
Character evaluation	0.13 (0.01)	6.90	<.001	0.08	0.04	0.93	1734.75	224
Empathic concern								
Act praise	0.05 (0.01)	4.88	<.001	0.10	0.05	0.90	1204.98	114
Character evaluation	0.20 (0.04)	8.65	<.001	0.09	0.04	0.92	1235.86	146

Note. *B* indicates the standardized β of the predicted pathway. *SE* = standard error; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; CFI = comparative fit index.

Table 6
Comparison Between Models With Alternative Paths in Study 3

	Distress		Concern	
Outcome variable	χ^2_{diff}	p	χ^2_{diff}	p
Pathway from anticipated affect to moral motives				
Act praise	4.16	.042	1.57	.220
Character evaluation	4.70	.030	1.34	.249
Pathway from anticipated affect to praise/character				
Act praise	41.98	<.001	8.49	.004
Character evaluation	0.156	.692	1.78	.182
Pathway from perceived affect to praise/character				
Act praise	44.09	<.001	9.63	.002
Character evaluation	2.07	.150	3.77	.052

Note. All model comparisons have $df_{diff} = 1$.

original model, though in all models, the original indirect path remained significant (indirect path estimates for the alternative models are shown in Supplemental Table S16). Generally, we observed the largest improvements in fit when adding paths from affect (anticipated or perceived) directly to praise judgments. This was not the case for character judgments, suggesting that, specifically for moral praise, some of the relationship between affect and judgments is not mediated by perceived moral motives. We return to this finding in the General Discussion section.

Discussion

The results of Study 3 replicated those of Study 2. However, unlike Study 2, Study 3 used a within-participants design where every participant rated each part of the proposed causal chain. We also measured perceived helper affect (which we had not measured in Study 2). As hypothesized, the more participants themselves expected to feel distress or empathic concern in a scenario, the more they thought someone who helped would feel these emotions. In turn, helpers who were perceived as more emotional were seen as possessing more moral motives, and higher moral motives were associated with greater moral praise for helping and more positive moral character judgments of the actor. In all four structural equation models we tested, the overall indirect effects for this path differed significantly from zero. As in Study 2, this was the case even when controlling for the perceived benefits of helping (although unlike in Study 2, in the present study, we controlled for the effect of perceived benefits on both perceived moral motives and on praise/character judgments).

Unlike in Studies 1 and 2, relationships were, if anything, stronger for empathic concern than for distress. We will return to this finding in the General Discussion section.

Study 4

In Study 4, we moved from measuring the proposed causal path to manipulating it. Our account holds that perceivers base their perceptions of a helper's emotional state on their own anticipated emotional reactions to the situation. These emotion perceptions then have downstream consequences for perceived moral motives and, ultimately, praise and character judgments. If this reasoning is correct, breaking the link between the perceiver's anticipated affect and perceptions of the helper's emotional state ought to eliminate (or at least attenuate) the relationship between anticipated affect and moral praise. In Study 4, we randomly assigned participants to see one of the 20 helping scenarios we used previously. However, in the present study, participants were also randomly assigned to one of two conditions. The control condition was identical to Study 3. In the experimental condition (henceforth referred to as the *unemotional* condition), we told participants that due to a brain condition, the helper relied only on rational calculation in deciding what to do. We predicted that in the unemotional condition, participants would be less likely to use their own anticipated affect to infer the target's emotionality, and that this would have downstream consequences for judgments of moral motives, praise, and character. We preregistered our planned sample size, predictions, and analyses at https://aspredicted.org/MYR_QMT. All analyses reported here can therefore be considered confirmatory, unless otherwise specified.

Method

Participants

We recruited a total of 1,002 U.S. adult participants from Prolific.co, a prescreened online labor market for behavioral research. Of these, 958 passed our preregistered attention check (which required participants to select a grammatically correct word that completes the sentence: "What time ____ bus leave?"). The final sample thus consisted of 958 participants total (M_{age} = 38.3, range = 18–93; 48.5% male, 48.7% female, 2.8% other), with 476 participants assigned to the unemotional condition, and 482 participants assigned to the control condition. Participants were paid £.50 (about $.65 U.S.) for their time.

Procedure

As in Study 3, participants first read one randomly chosen scenario of the 20 shown in Table 1. Each scenario asked the participant to imagine themselves in a situation where helping was possible (but did not describe a helping behavior). Participants then completed the same 14-item anticipated affect measure used in Study 3.

Next, participants read the same scenario but with the addition of a protagonist who helped (these are the same scenarios as were used in Studies 1–3; the full text of each is shown in Supplemental Table S3). Participants in the control condition read one of the scenarios exactly as in Study 3. Participants randomly assigned to the unemotional condition first read a description adapted from Critcher et al. (2020), which explained that the protagonist was born with a neurological disorder that caused them to be unemotional and rational:

> [Protagonist] is missing the part of their brain that allows them to have strong emotional impulses that signal what is morally right or wrong. Instead, all [protagonist] can do is use rational calculation to calculate what is the right thing to do. [Protagonist] was born that way: not having emotional reactions to others, but acting like a "computer" to compute what is right or wrong.

Immediately afterward, participants in the unemotional condition read the scenario describing the protagonist's helping behavior (from this point, these were identical to those in the control condition). Participants then completed (in random order) the measures of perceived helper emotions, moral motives, perceived benefits, and act/actor judgments from Study 3.

Table 7
Indirect Paths From Structural Equation Models in Study 4 (Control Group)

	Predicted pathway			Model fit				
Outcome variable	*B* (*SE*)	*Z*	*p*	RMSEA	SRMR	CFI	χ^2	*df*
Distress								
Act praise	0.02 (0.01)	2.35	.019	0.08	0.05	0.94	804.44	184
Character evaluation	0.09 (0.04)	3.78	<.001	0.08	0.04	0.94	869.45	224
Empathic concern								
Act praise	0.04 (0.02)	2.94	.003	0.10	0.05	0.91	655.51	114
Character evaluation	0.16 (0.05)	5.36	<.001	0.09	0.05	0.92	710.59	146

Note. *B* indicates the standardized β of the predicted pathway. *SE* = standard error; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; CFI = comparative fit index.

Results

Replication of Study 3

First, we replicated the analysis from Study 3 using only participants in the control group. Replicating our past findings, in all four models, the hypothesized indirect paths (from anticipated emotion to perceived emotion, to moral motives, and finally to praise/character judgments) were statistically significant. Fit statistics and indirect path tests for each model are shown in Table 7.[4] As in Study 3, indirect effects were about four times larger for moral character than for moral praise.

Effects of the Manipulation on Dependent Variables

Descriptive statistics by condition for each dependent variable (perceived helper emotional concern and distress, moral motives, praise, and helper moral character) are shown in Table 8. This table also shows tests of the between-condition differences in each of these variables. Differences between conditions were significant in the expected direction (i.e., lower helper emotionality, moral motives, and moral character in the unemotional condition) for every variable except praise, where the difference between conditions was in the expected direction but not significant ($p = .07$).[5]

Multigroup Measurement Model

Following a process similar to the steps outlined in Alley et al. (2023), we attempted to develop a measurement model that was invariant across conditions. This process, described in full in the Supplemental Materials, led us to develop a model with 22 of 34 items constrained across conditions (a list of these items is in Supplemental Table S19; model fit: $\chi^2 = 3367.47$, $df = 1{,}039$, CFI = 0.93, RMSEA = .08, SRMR = .06). We did not find evidence to conclude that this model was significantly different from an unconstrained model, where all items except one anchor item for each construct were allowed to freely vary (intercepts and loadings) across conditions ($\chi^2_{diff} = 24.6$, $df_{diff} = 15$, $p = .055$). Thus, this partially constrained model meets the standard for partial invariance across groups. Past research has found that when multiple-group models are partially constrained across groups, their parameter estimates for the relationships between factors, including regression coefficients, are similar to those produced by fully constrained models (Schmitt et al., 2011; Schmitt & Kuljanin, 2008). Thus, our partially constrained model with measurement invariance across conditions is sufficient to test our hypothesized pathway, and we proceeded with this model as our baseline model for subsequent analyses.

Path Analysis

Our main preregistered prediction was that participants in the unemotional condition would be less likely to use the pathway shown in Figure 3 to make inferences about the protagonist's actions and moral character. To test this, for each set of questions, we fit two multigroup SEM models (with condition as the grouping factor) to the complete data set. In the first model, we constrained each parameter in our key pathway to be identical across conditions. In the second model, we allowed each parameter in the key pathway to freely vary across conditions. Finally, we compared the two models to test whether allowing the parameters in the pathway to vary across conditions significantly improved model fit. For three of the four models, this was the case (see Table 9).

Relationship Between Anticipated and Perceived Helper Emotion

The test of between-condition differences in the full paths above is the most comprehensive test of whether the manipulation of helper emotionality affected the process by which participants made their judgments. However, because it tests differences in the entire path, some of which are of less theoretical interest, this test has lower specificity in that theoretically irrelevant path differences may cause the test to reject the null of no differences in paths between conditions. Thus, we also tested a more specific prediction: participants should be less likely to use their anticipated emotion to infer helper emotion in the unemotional condition. As with above, we tested this by comparing a model with this pathway constrained across conditions to a model where the pathway was allowed to vary across conditions.[6] For both empathic concern ($\chi^2_{diff} = 6.44$, $df = 1$, $p = .011$) and distress ($\chi^2_{diff} = 70.4$, $df = 1$, $p < .001$), the constrained

[4] We also tested indirect paths in the alternative models described above (indirect path estimates for the alternative models are shown in Supplemental Table S20). The original indirect paths remained statistically significant in most but not all of the alternative models. The difference from Study 3 (where all indirect paths remained significant) is likely due to the smaller sample size here because we are only using the control condition data ($n = 482$; vs. $N = 959$ in Study 3).

[5] Our preregistration did not include these simple tests of differences between conditions, but we include them here for completeness.

Table 8
Descriptive Statistics by Condition and Between-Condition Differences for Dependent Variables in Study 4

Measures	Control condition	Unemotional condition	Difference between conditions				
	M (*SD*)	*M* (*SD*)	$M_{control} - M_{emotional}$ (95% CI)	*t*	*df*	*p*	*d* (95% CI)
Helper emotional concern	5.53 (1.23)	2.48 (1.77)	3.05 [2.86, 3.25]	31.01	846.86	<.001	2.01 [1.85, 2.16]
Helper distress	3.38 (1.88)	2.06 (1.38)	1.32 [1.12, 1.53]	12.44	883.62	<.001	.80 [0.67, 0.93]
Moral motives	5.73 (1.13)	3.56 (1.86)	2.17 [1.97, 2.36]	21.77	781.49	<.001	1.41 [1.27, 1.55]
Praise	5.09 (1.20)	4.95 (1.25)	.14 [−.01, .30]	1.84	953.34	.07	.12 [−0.01, 0.25]
Moral character	6.03 (.93)	5.11 (1.23)	.92 [.78, 1.06]	13.06	883.77	<.001	.85 [0.71, 0.98]

Note. CI = confidence interval.

model was a worse fit for the data. Additionally, examining the size of the coefficients in the unconstrained model revealed that for both distress and concern, the relationship between anticipated and perceived helper emotion was weaker in the unemotional condition (distress: $\beta = 0.47$, $SE = 0.08$; concern: $\beta = 0.14$, $se = 0.09$) than in the control condition (distress: $\beta = 0.81$, $SE = 0.12$, concern: $\beta = 0.44$, $SE = 0.08$).

Discussion

The results of Study 4 supported the proposed causal account. Describing helpers as unable to feel emotion reduced the extent to which participants saw them as morally motivated and possessing positive moral character. In the unemotional condition, participants relied less on their own anticipated affective states to infer the actor's affective reactions, and tests of the overall fit of the proposed causal pathway (from anticipated affect, to perceived helper affect, to moral motives, to praise and character judgments) showed that these pathways differed significantly between conditions. In sum, these results suggest that interrupting the link between anticipated and perceived affect for prosocial others has significant downstream consequences for moral judgments.

General Discussion

In four preregistered studies, we investigated whether people use anticipated affect as a guide to praise and character judgments. Previous research suggests that prototypical altruistic behavior is seen as emotionally motivated, and that emotionally motivated altruists are seen as acting more out of genuine moral concern (Ames et al., 2004; Barasch et al., 2014; Montealegre et al., 2021). We therefore reasoned that helping in situations that arouse a stronger anticipated affective response would be praised more, and that helpers in affectively arousing situations would be seen as having more positive moral character. These hypotheses were supported in four studies, of which two (Studies 1 and 2) used between-participant designs and two (Studies 3 and 4) used within-participant designs. All studies used the same diverse set of naturalistic helping behaviors and scenarios.

Differences Between Distress and Empathic Concern

Studies 1 and 2 found weaker effects on praise and character judgments for anticipated empathic concern than for anticipated distress. However, in Studies 3 and 4, anticipated empathic concern was, if anything, more strongly predictive of praise and character judgments than anticipated distress. We believe these differing results between studies may result from design differences. Studies 1 and 2 were between-participants, with different groups of participants giving ratings of anticipated affect and praise/character judgments (thus, the hypothesis-testing analyses were on scenario-level aggregate ratings), while Studies 3 and 4 were within-participants, with every participant providing ratings for every item. When aggregating ratings to the scenario level (as we did in Studies 1 and 2), the level of agreement between raters becomes relevant, as lower agreement will produce less reliable aggregate scores (all else equal). We found that agreement between participants was particularly low for anticipated empathic concern. The results of Studies 3 and 4—in which anticipated empathic concern was strongly related to praise and character judgments—suggest that this low agreement is not random error (because random error would reduce effect sizes in a within-participants design as well). Rather, people seem to vary substantially in how much empathic concern they expect to feel in a given situation. The cause of this is an intriguing question for future research.

Past theorizing links empathic concern to "true" altruism (i.e., helping even when the helper had the option to exit the situation) and

Table 9
Comparison Between Constrained and Unconstrained Models in Study 4

Outcome	Unconstrained		Constrained		Difference		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	*p*
Distress							
Act praise	2155.55	422	2263.05	425	101	3	<.001
Character evaluation	1515.46	458	1627.27	461	89.2	3	<.001
Empathic concern							
Act praise	1854.22	272	1862.77	275	6.85	3	.077
Character evaluation	1207.80	300	1223.61	303	10.3	3	.013

links distress to merely contingent altruism (such that people no longer help when escape is possible; see Batson et al., 1983, 1987; for more empirical comparisons of empathic concern and distress, see FeldmanHall et al., 2015). However, perceivers seem to see empathic concern and distress as equally laudable reasons to help (Barasch et al., 2014). This suggests that lay people do not see these constructs in the same way researchers do. The current results reinforce this conclusion—both distress and empathic concern are related to greater praise for helpers and more positive judgments of their moral character.

We focused on distress and empathic concern because they have been studied extensively and are known to be morally relevant (Barasch et al., 2014; Batson et al., 1983, 1987). However, we do not believe the relationship between anticipated affect and moral judgment is necessarily limited to these specific states. In fact, even emotions such as disgust and anger can be morally relevant under the right circumstances—for example, as responses to perceived injustice (Chapman, 2018; Hutcherson & Gross, 2011). Thus, it is plausible that in some cases, people might rely on (for example) anticipated anger when judging how much someone should be praised for helping victims of unjust behavior. Of course, this is speculative, and the extent to which anticipation of other affective states affects moral judgments is a question for future research.

The Importance of Moral Motives

Inferences about motives (i.e., psychological states directed at some desired goal) play a large role in moral judgments, but these have been studied less by moral psychologists than act and character judgments (Carlson et al., 2022). A key part of our account is that observer inferences about moral motives mediate the relationship between anticipated affect and moral judgments. This is an important distinction between the current research and previous, related work. For example, Yudkin et al. (2019) examined the relationship between people's expectations of how they would feel when helping in different situations (what they called "action value") and perceptions of the praiseworthiness of helping. They found that action value ratings were associated with praiseworthiness ratings, even when controlling for perceived benefits to the targets of helping. The authors explained this relationship by arguing that positive affect (i.e., a "warm glow"; see Andreoni, 1990) often motivates helping behavior, and that expected affect likewise directly affects evaluations of helping. Thus, their account does not posit any role for inferences about the motivations of helpers (indeed, their scenarios never included descriptions of specific helpers but rather asked participants to rate the praiseworthiness of helping in the abstract). In contrast, our account focuses on the inferences observers make about people's motivations for helping, and we find that these seem to play a key role in praise and character judgments.

Past research has manipulated perceived motives directly by explicitly informing participants about the actor's motivation to help (e.g., Ames et al., 2004; Barasch et al., 2014; Montealegre et al., 2021). Although in some cases people might be exposed to this information in everyday life (e.g., because the helper communicates their motives to others); in many cases, they are not. In cases where information about motivation is lacking, perceivers must infer it from characteristics of the behavior or the situation (Ames & Johar, 2009; Critcher et al., 2013). Indeed, there is evidence that people do sometimes infer morally relevant mental states from aspects of the situation (such as what is visually salient to the agent), and that this exerts downstream effects on moral judgments (Critcher et al., 2020). The present studies apply this logic to inferences about emotional motivation (or lack of it). When people need to predict how others would respond to emotionally arousing events, they ask themselves how they themselves would feel (Van Boven & Loewenstein, 2003). Thus, when people encounter a situation in which someone acted prosocially, they imagine how they themselves would have felt in that situation, and then use that anticipated experience to infer how the actor must have felt. To the extent that these inferred states match the motivations that people expect for prosocial behavior (i.e., feeling emotionally moved), the helping behavior and the helper are viewed more positively.

This seems to be a very robust phenomenon. Although we were able to reduce the extent to which participants used their own emotional responses to infer helpers' emotions, even the strong manipulation we used did not eliminate these inferences entirely (i.e., the relationship between anticipated and perceived affect "unemotional" condition was still positive for both empathic concern and distress). In another study reported in detail in the Supplemental Material, we found that a weaker manipulation of helper affect (i.e., describing the helper as "calm and unemotional") did not moderate the use of anticipated affect at all. Compared to Study 4, where the actor's reason for lacking affect is made clear to participants, Supplemental Study S2 left more ambiguity about the reason for the actor's lack of emotion. Evidently, any such ambiguity makes it more likely that observers will use their anticipated affect to make moral judgments, even when it would objectively seem unwarranted. Thus, it appears that the use of one's own anticipated emotions to predict those of others is a strong default (consistent with the central role of projection in many social judgments; see Allport, 1924; Krueger & Clement, 1994; Robbins & Krueger, 2005; Ross et al., 1977).

Assumptions, Limitations, and Future Directions

Causal Inference Assumptions

Our account is causal: We propose that differences in anticipated affect cause differences in downstream judgments. Studies 2–4 tested this account in different ways. In Studies 2 (between participants) and 3 (within participants), we used designs in which we measured but did not manipulate the hypothesized causal chain (i.e., "measurement-of-mediation" designs). Measurement-only designs can certainly disconfirm a causal account (e.g., if the measured variables do not relate to each other as hypothesized; see Fiedler et al., 2018), but how strongly they support a causal account depends on how effectively researchers are able to rule out possible confounding variables throughout the causal chain (Rohrer et al., 2022). Our analyses controlled for perceived benefits to the target of helping, which could plausibly affect all elements of the causal chain (i.e., perceiving someone to be greatly in need of help could cause heightened affect, stronger perceived moral motives in the helper,

[6] In our preregistration, we stated that we would test this hypothesis using linear regression. However, an anonymous reviewer pointed out that comparing the significance of paths across conditions is a more appropriate test. Thus, we report this test here. The results of the preregistered tests (which supported our hypotheses) are reported in the Supplemental Materials.

and more positive praise/character judgments). However, other unmeasured variables might be related to at least two parts of the causal chain. For example, perceived cost to the helper might be related to moral motives, praise, and character judgments (similar to how expected effort can make prosocial acts more attractive; see Olivola & Shafir, 2013). Likewise, global positive or negative impressions of the helper could plausibly be related to moral motives, praise, and character judgments. Given the size of the indirect effects in Studies 3 and 4 (i.e., the better-powered within-participants studies), we are skeptical that these or other unmeasured variables would entirely account for the observed effects. However, the assumption that all relevant confounding variables have been controlled is a very strong one (albeit one that is required in the typical mediation analysis; see Rohrer et al., 2022).

In Study 4, we therefore combined measurement with a "moderation-of-process" design in which we experimentally manipulated perceivers' reliance on their own anticipated affect to infer helpers' emotional states, which is a key part of the hypothesized causal chain. This design is not susceptible to the confounding-variable concerns that apply to measurement-of-mediation (for more on the trade-offs between different designs for process evidence, see Spencer et al., 2005). The results of this approach largely supported the proposed causal account. The unemotional condition elicited significantly lower helper emotionality, moral motives, and moral character compared to the control condition, with the exception of praise, where the difference was directionally consistent but not statistically significant ($p = .07$). Model comparisons showed that for all models except empathic concern/praise, the manipulation significantly affected the proposed causal pathway.

Alternative Causal Models

In Study 3, we tested the plausibility of alternative models by testing whether adding additional paths between variables improved model fit. We found that the largest improvements resulted from the addition of direct paths between anticipated and perceived affect and moral praise. In contrast, adding these paths for character judgments did not significantly improve model fit. This is consistent with a direct relationship between emotion and moral praise (in addition to the indirect relationship via perceived moral motives). This may be because, as Yudkin et al. (2019) suggested, there is a direct relationship between affect and moral praise (e.g., because scenario-evoked affect cues praise apart from any inferences about the helper). It could also be that there is an indirect relationship via a variable that we did not measure in these studies. In that case, though, it is not clear why this unmeasured path would not also affect character judgments (where we found no fit improvements from adding additional paths to the model). We think investigating these questions is an important topic for future research.

Limitations

One limitation of the current research is the lack of demographic variables collected in these online samples. To reduce items and increase task brevity, we collected only age, gender, and English fluency across study samples (see Table 10, for full description of limitations). We recommend that future studies investigate the relationship between anticipated affect and praise judgments in non-Western, Educated, Industrialized, Rich and Democratic and majority non-White populations.

In both the between- and within-participant studies, anticipated affect ratings were in response to scenarios that did not include a helping behavior (i.e., they described only the situation where helping was possible, not a helper's actions). In the between-participant studies, this was done by asking different groups of participants to provide ratings of anticipated affect and helpers; in the within-participant studies, participants first rated their anticipated affect for a scenario and then their evaluations of an actor who helped in the same scenario. We chose to design the studies this way for two reasons. First, we worried that including a helping behavior when soliciting affect ratings might artifactually produce the predicted results (e.g., because participants might give affect ratings consistent with their evaluations of the helper they had just read about). Second, we thought that including a helping behavior risked confusing participants about what they were supposed to be rating, and that many would tell us their anticipated affect in response to the helping behavior rather than the eliciting situation. The downside of this approach, however, is that the scenarios in which participants evaluated helpers are not identical to the ones in which they predicted their own affect. To the extent that including a helper might (nonartifactually) change anticipated affect, this would work against detecting a relationship between anticipated affect and helper evaluations (i.e., it would bias the effects estimated here downward). It is not clear, though, whether this should be considered a limitation, as we believe that all things considered, it is preferable to the alternative approach (i.e., including a helping behavior for anticipated affect ratings).

Table 10
Limitations Across Studies

Limitation	Study			
	Study 1	Study 2	Study 3	Study 4
Racial/ethnic demographic variables not collected[a]		X	X	X
No process measures	X			
Lower-powered between-participants design	X	X		X
Measurement-only design	X	X	X	

[a] We collected data on race and ethnicity for two waves of data collection, both of which are included in full detail in Supplemental Table S1, and found that participants were 72.8% White for the first wave of data collection (anticipated affect ratings) and 77.8% White for Study 1.

Conclusions and Implications

The current findings are congruent with a broader conclusion from research on prosociality. When evaluating prosocial acts, people are sensitive not only to the good brought about but also to the actor's motivations (Carlson et al., 2022). For many, the prototypical prosocial act is the result of feelings rather than dispassionate analysis (Berman et al., 2018; Montealegre et al., 2021). This may help explain resistance to the idea of effective altruism, which holds that donors ought to allocate their resources to maximize the welfare benefit per dollar given (Law et al., 2021; MacAskill, 2015). This cold-blooded focus on cost-effectiveness may, for many, feel incompatible with the emotional responses to need that people expect to motivate prosociality. In fact, advocates of effective altruism might be more persuasive if they emphasized the feelings underlying the desire to do the best possible over the calculations that determine the most effective way to allocate one's charitable spending.

Statement of Limitations

The main limitations of this research are the ethnic and racial homogeneity of these samples and the lack of demographic variables collected across the four studies. We collected data on race and ethnicity for two waves of data collection, both of which are included in full detail in Supplemental Table S1, and found that participants were 72.8% White for the first wave of data collection (anticipated affect ratings) and 77.8% White for Study 1. We recommend that future studies investigate the relationship between anticipated affect and praise judgments in non-Western, Educated, Industrialized, Rich and Democratic and majority non-White populations in order to extend the generalizability of these findings.

References

Alley, L. J., Axt, J., & Flake, J. K. (2023). *Convenience samples and measurement equivalence in replication research* [Unpublished manuscript]. Psychology Department, McGill University. https://osf.io/32unb

Allport, F. (1924). Social attitudes and social consciousness. *Social psychology* (pp. 320–335). Houghton Mifflin Company.

Ames, D. R., Flynn, F. J., & Weber, E. U. (2004). It's the thought that counts: On perceiving how helpers decide to lend a hand. *Personality and Social Psychology Bulletin*, *30*(4), 461–474. https://doi.org/10.1177/0146167203261890

Ames, D. R., & Johar, G. V. (2009). I'll know what you're like when i see how you feel: How and when affective displays influence behavior-based impressions. *Psychological Science*, *20*(5), 586–593. https://doi.org/10.1111/j.1467-9280.2009.02330.x

Anderson, R. A., Crockett, M. J., & Pizarro, D. A. (2020). A theory of moral praise. *Trends in Cognitive Sciences*, *24*(9), 694–703. https://doi.org/10.1016/j.tics.2020.06.008

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal*, *100*(401), 464–477. https://doi.org/10.2307/2234133

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), 393–413. https://doi.org/10.1037/a0037207

Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? *Advances in Experimental Social Psychology*, *20*, 65–122. https://doi.org/10.1016/S0065-2601(08)60412-8

Batson, C. D. (2010). Empathy-induced altruistic motivation. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature* (pp. 15–34). American Psychological Association. https://doi.org/10.1037/12061-001

Batson, C. D., Ahmad, N., Lishner, D. A., & Tsang, J.-A. (2002). Empathy and altruism. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 485–498). Oxford University Press.

Batson, C. D., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, *45*(3), 706–718. https://doi.org/10.1037/0022-3514.45.3.706

Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological Science*, *29*(5), 834–844. https://doi.org/10.1177/0956797617747648

Carlson, R. W., Bigman, Y. E., Gray, K., Ferguson, M. J., & Crockett, M. J. (2022). How inferred motives shape moral judgements. *Nature Reviews Psychology*, *1*(8), 468–478. https://doi.org/10.1038/s44159-022-00071-x

Chapman, H. A. (2018). A component process model of disgust, anger, and moral judgment. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 70–80). Guilford Press.

Critcher, C. R., Helzer, E. G., & Tannenbaum, D. (2020). Moral character evaluation: Testing another's moral-cognitive machinery. *Journal of Experimental Social Psychology*, *87*, Article 103906. https://doi.org/10.1016/j.jesp.2019.103906

Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science*, *4*(3), 308–315. https://doi.org/10.1177/1948550612457688

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. https://doi.org/10.1016/j.cognition.2008.03.006

Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: The aversion to harmful action. *Emotion*, *12*(1), 2–7. https://doi.org/10.1037/a0025071

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage*, *105*, 347–356. https://doi.org/10.1016/j.neuroimage.2014.10.043

Fiedler, K., Harris, C., & Schott, M. (2018). Unwarranted inferences from statistical mediation tests—An analysis of articles published in 2015. *Journal of Experimental Social Psychology*, *75*, 95–102. https://doi.org/10.1016/j.jesp.2017.11.008

Guglielmo, S., & Malle, B. F. (2019). Asymmetric morality: Blame is more differentiated and more extreme than praise. *PLOS ONE*, *14*(3), Article e0213544. https://doi.org/10.1371/journal.pone.0213544

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*(6202), 1340–1343. https://doi.org/10.1126/science.1251560

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social-functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*(4), 719–737. https://doi.org/10.1037/a0022408

Kahneman, D., & Ritov, I. (1994). Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty*, *9*(1), 5–37. https://doi.org/10.1007/BF01073401

Krueger, J. (1998). On the perception of social consensus. In *Advances in experimental social psychology* (Vol. 30, pp. 163–240). Academic Press. https://doi.org/10.1016/S0065-2601(08)60384-6

Krueger, J. (2000). The projective perception of the social world: A building block of social comparison processes. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 323–351). Springer US. https://doi.org/10.1007/978-1-4615-4237-7_16

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*(4), 596–610. https://doi.org/10.1037/0022-3514.67.4.596

Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, *36*(2), 249–277. https://doi.org/10.1207/S15327906MBR3602_06

Law, K. F., Campbell, D., & Gaesser, B. (2021). Biased benevolence: The perceived morality of effective altruism across social distance. *Personality and Social Psychology Bulletin*. *48*(3), 426–444. https://doi.org/10.1177/01461672211002773

Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation—A discussion and demonstration of basic features. *PLOS ONE*, *14*(7), Article e0219854. https://doi.org/10.1371/journal.pone.0219854

MacAskill, W. (2015). *Doing good better: Effective altruism and a radical new way to make a difference*. Guardian Faber Publishing.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

MacKinnon, D. P., & Pirlott, A. G. (2015). Statistical approaches for enhancing causal interpretation of the M to Y relation in mediation analysis. *Personality and Social Psychology Review*, *19*(1), 30–43. https://doi.org/10.1177/1088868314542878

McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science*, *31*(3), 227–242. https://doi.org/10.1177/0956797619900321

Miller, R., & Cushman, F. (2013). Aversive for me, wrong for you: First-person behavioral aversions underlie the moral condemnation of harm. *Social and Personality Psychology Compass*, *7*(10), 707–718. https://doi.org/10.1111/spc3.12066

Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*, *14*(3), 573–587. https://doi.org/10.1037/a0035361

Montealegre, A., Bush, L., Moss, D., Pizarro, D., & Jimenez-Leal, W. (2021, November 23). *Does maximizing good make people look bad?* PsyArXiv. https://doi.org/10.31234/osf.io/2zbax

Morrison, T. G., Morrison, M. A., & McCutcheon, J. M. (2017). Best practice recommendations for using structural equation modelling in psychological research. *Psychology*, *8*(9), 1326–1341. https://doi.org/10.4236/psych.2017.89086

Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of Behavioral Decision Making*, *26*(1), 91–105. https://doi.org/10.1002/bdm.767

Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science*, *14*(3), 267–272. https://doi.org/10.1111/1467-9280.03433

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209–233. https://doi.org/10.1037/a0020141

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3

Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2022). That's a lot to process! Pitfalls of popular path models. *Advances in Methods and Practices in Psychological Science*, *5*(2). https://doi.org/10.1177/25152459221095827

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Schmitt, N., Golubovich, J., & Leong, F. T. (2011). Impact of measurement invariance on construct correlations, mean differences, and relations with external correlates: An illustrative example using Big Five and RIASEC measures. *Assessment*, *18*(4), 412–427. https://doi.org/10.1177/1073191110373223

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, *18*(4), 210–222. https://doi.org/10.1016/j.hrmr.2008.03.003

Schumacker, R. E., & Lomax, R. G. (2010). *A beginners guide to structural equation modeling*. Routledge.

Shaw, L. L., Batson, C. D., & Todd, R. M. (1994). Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences. *Journal of Personality and Social Psychology*, *67*(5), 879–887. https://doi.org/10.1037/0022-3514.67.5.879

Siegel, J. Z., Crockett, M. J., & Dolan, R. J. (2017). Inferences about moral character moderate the impact of consequences on blame and praise. *Cognition*, *167*, 201–211. https://doi.org/10.1016/j.cognition.2017.05.004

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*(6), 845–851. https://doi.org/10.1037/0022-3514.89.6.845

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*(1), 72–81. https://doi.org/10.1177/1745691614556679

Uhlmann, E. L., Zhu, L. L., & Tannenbaum, D. (2013). When it takes a bad person to do the right thing. *Cognition*, *126*(2), 326–334. https://doi.org/10.1016/j.cognition.2012.10.005

Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin*, *29*(9), 1159–1168. https://doi.org/10.1177/0146167203254597

Van Boven, L., & Loewenstein, G. (2005). Empathy gaps in emotional perspective taking. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 284–297). Guilford Press.

Van Boven, L., Loewenstein, G., Dunning, D., & Nordgren, L. F. (2013). Changing places. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 48, pp. 117–171). Elsevier. https://doi.org/10.1016/B978-0-12-407188-9.00003-X

Yudkin, D. A., Prosser, A. M. B., & Crockett, M. J. (2019). Actions speak louder than outcomes in judgments of prosocial behavior. *Emotion*, *19*(7), 1138–1147. https://doi.org/10.1037/emo0000514

Received April 13, 2023
Revision received October 25, 2023
Accepted November 9, 2023 ■