Will I Be Negatively Evaluated When Help Fails? Prediction Bias in Well-Intentioned but Counter-productive Assistance

Authors: Shang Xuesong, Chen Zhuo, Lu Jingyi

Date: 2020-11-14T00:00:00+00:00

Abstract

Instances of well-intentioned assistance backfiring occur with appreciable frequency. Helpers who fail in their efforts often anticipate that recipients will harshly criticize them, potentially rendering them reluctant to offer help in the future. Is this prediction accurate? Across six studies (N = 1763), comparing helpers' predictions of recipients' reactions with recipients' actual responses, we identified a predictive bias among helpers when their well-intentioned efforts backfire: they overestimated the negativity of recipients' reactions. Conversely, when assistance proved successful, helpers exhibited either no predictive bias or a substantially attenuated one. This predictive bias arises because helpers focus predominantly on their own competence, assuming that recipients likewise attend to their competence, whereas recipients are primarily concerned with the helper' s warmth.

Full Text

Will I Be Negatively Evaluated After a Failed Attempt to Help? Prediction Bias in Well-Intentioned but Counter-productive Helping

Shang Xuesong, Chen Zhuo, Lu Jingyi
School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China

Abstract

Incidents of well-intentioned but counterproductive helping occur frequently. Helpers who fail in their attempts often believe that help recipients will harshly

criticize them, which may make them reluctant to offer help again. Is this prediction accurate? Through six studies (N = 1,763), comparing helpers' predictions of help recipients' reactions with help recipients' actual reactions, we identified a prediction bias in well-intentioned but counterproductive helping: helpers overestimated help recipients' negative reactions. When help was successful, helpers showed no prediction bias or a weaker bias. The cause of this prediction bias is that helpers focus more on their own competence, assuming that help recipients also focus on their competence, whereas help recipients actually care more about helpers' warmth.

Keywords: prediction bias, helping, competence, warmth, judgment and decision making

1.1 Prediction Bias in Helping Contexts

In social interactions, people' s predictions of others' thoughts and behaviors often deviate from others' actual situations. This type of prediction bias (misprediction) is ubiquitous. For example, people overestimate the likelihood that others will choose common options (Reit & Critcher, 2020); overestimate others' willingness to befriend high-status individuals (Garcia et al., 2019); expressers of gratitude underestimate recipients' positive feelings and overestimate their embarrassment (Kumar & Epley, 2018); individuals who share their achievements with others underestimate listeners' negative experiences (Scopelliti et al., 2015); speakers underestimate how much listeners like them (Boothby et al., 2018); storytellers predict that audiences prefer novel stories when in fact audiences prefer familiar ones (Cooney et al., 2017).

Prediction bias also frequently occurs in helping contexts (Wang & Xie, 2019). Before helping occurs, help-seekers underestimate the likelihood that others will agree to help (Flynn & Bohns, 2008), while helpers underestimate the embarrassment help-seekers feel when asking for help (Bohns & Flynn, 2010). During the helping process, help recipients underestimate the time and effort helpers are willing to expend (Newark et al., 2017); helpers tend to provide paternalistic help, neglecting help recipients' autonomy, whereas help recipients desire agentic help and crave autonomy (Schroeder et al., 2017). After helping is completed, helpers incorrectly predict the rewards help recipients are willing to give, assuming that help recipients will reciprocate based on the costs helpers incurred, when in fact help recipients reciprocate based on the benefits they received (Zhang & Epley, 2009).

Although researchers have examined prediction bias in helping contexts, most have focused on successful helping while neglecting failed attempts. Compared with successful helping, do helpers exhibit unique prediction bias when predicting help recipients' reactions after failed attempts? This question has important theoretical and practical significance. At the theoretical level, examining unique prediction bias in failed helping contexts can advance research on prediction bias and reveal the psychological mechanisms underlying it. At the practical level,

understanding this bias can reduce helpers' concerns after failure and encourage them to continue providing help, thereby promoting behaviors that benefit others and society.

1.2 Warmth and Competence in Well-Intentioned but Counterproductive Helping

Helping behavior involves two dimensions: motivation and outcome. Regarding motivation, as long as it is genuine help, the helper' s actions should be well-intentioned regardless of the outcome. Therefore, this research only examines cases where helpers have good intentions; situations where individuals harm others under the guise of helping are beyond the scope of this paper. Regarding outcome, successful helping benefits help recipients and demonstrates the helper' s strong competence. Conversely, failed helping causes losses to help recipients, which is the unique characteristic of well-intentioned but counterproductive helping. In such cases, the failed outcome reflects the helper' s lack of competence.

In summary, well-intentioned but counterproductive helping refers to situations where helpers offer assistance to those in difficulty with good intentions, but the attempt fails and help recipients suffer losses as a result. For example, a helper offers to park a car but causes a scratch, or helps move a suitcase but damages valuable items inside. In these situations, helpers demonstrate warmth while simultaneously exposing their lack of competence.

Warmth and competence are the two fundamental dimensions of social cognition through which people judge others (Fiske et al., 2007). Warmth reflects a person' s intentions; friendliness, goodwill, sincerity, and trustworthiness all embody warmth. Competence reflects whether a person can realize their intentions; intelligence, capability, creativity, and efficiency all embody competence (Fiske et al., 2002).

Although both warmth and competence are important for social judgment, people value others' warmth more than others' competence when evaluating others, whereas they value their own competence more than their own warmth when evaluating themselves (Wojciszke, 2005). Theoretically, warmth is an other-profitable trait, so people care about whether others are warm because they can benefit from others' warmth (e.g., receiving help). Competence is a self-profitable trait, so people care about whether they themselves are competent because they can benefit from their own competence (e.g., achieving career success). Empirical research supports this view. Wojciszke (1994) found that participants' evaluations of others were more influenced by others' warmth and less by others' competence, whereas evaluations of themselves were more influenced by their own competence and less by their own warmth. Abele and Wojciszke (2007) asked participants to choose a training program for themselves or for others. Results showed that participants tended to choose warmth-enhancing training for others and competence-enhancing training for themselves. The dif-

ferential weighting of competence and warmth when judging self versus others leads expressers of gratitude to focus more on whether their wording is appropriate, whereas recipients of gratitude focus more on the positive social meaning embedded in the gratitude expression (Kumar & Epley, 2018).

Based on this logic, in well-intentioned but counterproductive helping contexts, help recipients should value warmth more and competence less when judging helpers, compared with helpers' predictions about help recipients.

1.3 Anchoring Effect and Prediction Bias

Because people cannot access others' thoughts but can easily retrieve their own, when estimating others' thoughts, people first anchor on their own perspective and then adjust toward others' perspective. However, this adjustment is often insufficient (Tamir & Mitchell, 2013), leading to prediction bias such as overestimating the consistency between others' attitudes and one' s own (Krueger & Clement, 1994).

In well-intentioned but counterproductive helping contexts, helpers value their own competence highly. Due to the anchoring effect, they incorrectly predict that help recipients also care about their competence. In reality, help recipients care more about helpers' warmth. Because helping demonstrates warmth while failed helping demonstrates low competence, helpers' predictions of help recipients' reactions are more negative than help recipients' actual reactions.

We propose the following hypotheses:

Hypothesis 1: In well-intentioned but counterproductive helping contexts, helpers will overestimate help recipients' negative reactions (i.e., underestimate their positive reactions). In our studies, we compare helpers' predictions of help recipients' reactions with help recipients' actual reactions to test whether helpers exhibit prediction bias. Studies 1, 2, 4, and 5 employ four indicators of help recipients' reactions: gratitude toward helpers, satisfaction with helpers, willingness to seek help from the same helper again, and willingness to recommend the helper to others. The first two indicators reflect immediate experiences and typical psychological feelings, while the latter two reflect future behavioral intentions. Study 3 uses a monetary indicator related to interests to more comprehensively examine help recipients' reactions.

Hypothesis 2: The reason for the above prediction bias is that, compared with helpers' predictions, help recipients actually pay more attention to helpers' warmth and less attention to helpers' competence.

1.4 Competing Hypotheses

In addition to the competence-warmth hypothesis, two other hypotheses might seemingly explain helpers' prediction bias in well-intentioned but counterproductive helping.

The first is the **spotlight effect**, which refers to people' s tendency to overestimate the degree to which others notice them (Gilovich et al., 2000). People incorrectly believe that others pay as much attention to them as they do themselves, as if they are standing under a spotlight. In reality, others do not pay that much attention. According to the spotlight effect, people exaggerate the impact of their own behavior on others. In well-intentioned but counterproductive helping, due to the spotlight effect, helpers may focus excessively on their own failure, amplify its impact on help recipients, and consequently overestimate help recipients' negative reactions. However, according to the spotlight effect, people amplify all their behaviors, both successful and unsuccessful. Therefore, helpers should overestimate both help recipients' negative reactions after failure and positive reactions after success. Under this explanation, the direction of prediction bias should differ between success and failure conditions. If no prediction bias exists in the success condition, or if the bias direction is the same as in the failure condition, the spotlight effect explanation can be ruled out.

The second is **social desirability**, which refers to people' s tendency to act in socially approved ways (Zerbe & Paulhus, 1987). In helping contexts, social norms require help recipients to reciprocate generously. Therefore, after receiving help, even if the outcome is unsatisfactory, help recipients may conceal their dissatisfaction and appear more grateful than they actually are, causing helpers to overestimate help recipients' negative reactions (i.e., underestimate their positive reactions). Given social norms, help recipients should exaggerate their gratitude regardless of whether the help succeeded or failed. Therefore, according to the social desirability explanation, helpers should underestimate help recipients' positive reactions to the same degree in both success and failure conditions. If no prediction bias or only a weak bias exists in the success condition, the social desirability explanation can be ruled out.

In contrast, according to the hypothesis that "helpers focus on competence while help recipients focus on warmth," in the success condition, because helpers demonstrate both warmth and competence, helpers' predictions about help recipients should show little bias. In the failure condition, helpers should exhibit prediction bias. In other words, the prediction bias in the failure condition should be greater than that in the success condition.

To distinguish among these three hypotheses, this research includes both success and failure conditions. By comparing the direction and magnitude of prediction bias across these conditions, we can test the reasons for helpers' prediction bias in well-intentioned but counterproductive helping and reveal what makes failure contexts unique compared with success contexts.

1.5 Overview of Studies

We conducted six studies to test helpers' prediction bias in well-intentioned but counterproductive helping and reveal its underlying cause. Study 1 used a roller-skating teaching scenario with success and failure conditions to test whether

prediction bias exists in failed helping, whether it differs in direction from that in successful helping, and whether its magnitude is greater than in successful helping, thereby ruling out competing hypotheses and revealing the uniqueness of failure contexts. Studies 2a and 2b used a car parking scenario to examine whether the unique prediction bias in well-intentioned but counterproductive helping remains stable across proactive and reactive helping. Study 3 used a guest-serving scenario to examine whether prediction bias in well-intentioned but counterproductive helping can affect subsequent monetary behaviors. Study 4 used a box-moving scenario to test the stability of prediction bias and examine whether helpers' and help recipients' perceptions of helpers' competence and warmth cause the prediction bias. Study 5 used a box-moving scenario to further explore the cause of prediction bias by recording participants' real-time thoughts during prediction or judgment.

According to G*Power calculations, for a medium effect size ($f = 0.25$) and statistical power of 0.95, at least 52.5 participants per condition are required. In all our experiments, we recruited 60 to 70 participants per condition to ensure adequate statistical power.

2 Study 1: Teaching Roller Skating

Study 1 used a roller-skating teaching scenario to test Hypothesis 1 and examine whether helpers exhibit prediction bias about help recipients' reactions in well-intentioned but counterproductive helping contexts. We included both success and failure conditions to test whether prediction bias in failed helping is unique compared with successful helping and to examine whether the spotlight effect and social desirability bias can adequately explain prediction bias in well-intentioned but counterproductive helping.

2.1 Participants

Two hundred eighty-one participants (138 males, 143 females) with a mean age of 27.31 years ($SD = 7.13$) completed the study. The study employed a 2 (outcome: success/failure) $\times$ 2 (role: helper/help recipient) between-subjects design.

2.2 Procedure

Participants in the helper condition read a scenario imagining themselves at a park with their friend Xiao Zhang, where many people were roller skating. Xiao Zhang, who could not skate, was interested and asked the participant, who could skate, to teach him. The participant agreed. In the success condition, participants were told that during the process, they were very careful, Xiao Zhang skated smoothly, and no falls or injuries occurred. In the failure condition, participants were told that during the process, they accidentally fell, causing Xiao Zhang to fall and suffer several scrapes.

Participants in the help recipient condition read a similar scenario but imagined themselves as Xiao Zhang being taught by their friend. After reading the scenario, helpers predicted help recipients' reactions: "Do you think Xiao Zhang would blame or be grateful to you?" (-7 = very blameful, 7 = very grateful), "How satisfied do you think Xiao Zhang would be with you?" (-7 = very dissatisfied, 7 = very satisfied), "If Xiao Zhang encountered difficulties in the future, do you think he would be willing to ask you for help again?" (1 = not at all willing, 7 = very willing), and "If others wanted to ask you for help in the future, how do you think Xiao Zhang would react?" (1 = strongly discourage, 7 = strongly support).

Help recipients rated their own reactions: "Would you blame or be grateful to Xiao Zhang?"(-7 = very blameful, 7 = very grateful), "How satisfied are you with Xiao Zhang?" (-7 = very dissatisfied, 7 = very satisfied), "If you encountered difficulties in the future, would you be willing to ask Xiao Zhang for help again?" (1 = not at all willing, 7 = very willing), and "If others wanted to ask Xiao Zhang for help in the future, how would you react?" (1 = strongly discourage, 7 = strongly support). Finally, participants reported their gender and age.

2.3 Results and Discussion

First, we conducted a 2 (outcome) × 2 (role) ANOVA on gratitude scores (Table 1). The main effect of outcome was significant, $F(1, 277) = 346.37$, $p < 0.001$, $p^2 = 0.56$. Participants predicted or rated higher gratitude in the success condition ($M = 5.76$, $SD = 1.40$) than in the failure condition ($M = 0.94$, $SD = 2.94$). The main effect of role was significant, $F(1, 277) = 30.31$, $p < 0.001$, $p^2 = 0.10$. Help recipients' actual gratitude ($M = 4.04$, $SD = 2.89$) was higher than helpers' predictions ($M = 2.62$, $SD = 3.61$).

Most importantly, the outcome × role interaction was significant, $F(1, 277) = 7.84$, $p = 0.005$, $p^2 = 0.03$. In the success condition, helpers underestimated help recipients' gratitude, $F(1, 137) = 9.23$, $p = 0.003$, $p^2 = 0.06$. In the failure condition, helpers underestimated help recipients' gratitude to a greater extent, $F(1, 140) = 21.86$, $p < 0.001$, $p^2 = 0.14$. This result indicates that prediction bias was greater in well-intentioned but counterproductive helping than in successful helping.

Second, we conducted a 2 (outcome) × 2 (role) ANOVA on satisfaction scores (Table 1). The main effect of outcome was significant, $F(1, 277) = 338.49$, $p < 0.001$, $p^2 = 0.55$. Participants predicted or rated higher satisfaction in the success condition ($M = 5.78$, $SD = 1.36$) than in the failure condition ($M = 0.83$, $SD = 3.17$). The main effect of role was significant, $F(1, 277) = 41.76$, $p < 0.001$, $p^2 = 0.13$. Help recipients' actual satisfaction ($M = 4.15$, $SD = 2.90$) was higher than helpers' predictions ($M = 2.42$, $SD = 3.80$).

Most importantly, the outcome × role interaction was significant, $F(1, 277) = 8.69$, $p = 0.003$, $p^2 = 0.03$. In the success condition, helpers underestimated help recipients' satisfaction, $F(1, 137) = 19.18$, $p < 0.001$, $p^2 = 0.12$. In the

failure condition, helpers underestimated help recipients' satisfaction to a greater extent, $F(1, 140) = 26.85$, $p < 0.001$, $p^2 = 0.16$. This result indicates that prediction bias was greater in well-intentioned but counterproductive helping than in successful helping.

Third, we conducted a 2 (outcome) × 2 (role) ANOVA on willingness to seek help again (Table 1). The main effect of outcome was significant, $F(1, 277) = 109.02$, $p < 0.001$, $p^2 = 0.28$. Participants predicted or rated higher willingness in the success condition ($M = 6.24$, $SD = 0.72$) than in the failure condition ($M = 4.99$, $SD = 1.28$). The main effect of role was significant, $F(1, 277) = 16.68$, $p < 0.001$, $p^2 = 0.06$. Help recipients' actual willingness ($M = 5.86$, $SD = 1.15$) was higher than helpers' predictions ($M = 5.37$, $SD = 1.23$).

Most importantly, the outcome × role interaction was significant, $F(1, 277) = 5.64$, $p = 0.018$, $p^2 = 0.02$. In the success condition, helpers accurately predicted help recipients' willingness to seek help again, $F(1, 137) = 2.85$, $p = 0.094$, $p^2 = 0.02$. In the failure condition, helpers underestimated help recipients' willingness, $F(1, 140) = 14.22$, $p < 0.001$, $p^2 = 0.09$. This result indicates that prediction bias existed only in the failure condition—helpers in well-intentioned but counterproductive helping underestimated help recipients' willingness to seek help again.

Fourth, we conducted a 2 (outcome) × 2 (role) ANOVA on recommendation willingness (Table 1). The main effect of outcome was significant, $F(1, 277) = 76.42$, $p < 0.001$, $p^2 = 0.22$. Participants predicted or rated higher recommendation willingness in the success condition ($M = 5.91$, $SD = 0.93$) than in the failure condition ($M = 4.84$, $SD = 1.15$). The main effect of role was significant, $F(1, 277) = 7.64$, $p = 0.006$, $p^2 = 0.03$. Help recipients' actual recommendation willingness ($M = 5.54$, $SD = 1.08$) was higher than helpers' predictions ($M = 5.20$, $SD = 1.24$).

Most importantly, the outcome × role interaction was significant, $F(1, 277) = 5.86$, $p = 0.016$, $p^2 = 0.02$. In the success condition, helpers accurately predicted help recipients' recommendation willingness, $F(1, 137) = 0.07$, $p = 0.792$, $p^2 < 0.01$. In the failure condition, helpers underestimated help recipients' recommendation willingness, $F(1, 140) = 11.67$, $p < 0.001$, $p^2 = 0.08$. This result indicates that prediction bias existed only in the failure condition—helpers in well-intentioned but counterproductive helping underestimated help recipients' willingness to recommend them.

Study 1 supported Hypothesis 1: In well-intentioned but counterproductive helping, helpers' predictions of help recipients' reactions were more negative than help recipients' actual reactions. In both success and failure conditions, helpers' prediction scores were lower than help recipients' actual ratings, ruling out the spotlight effect explanation. In the success condition, helpers underestimated help recipients' gratitude and satisfaction, which may be due to help recipients' social desirability—help recipients may hide their dissatisfaction due to social norms. Most importantly, helpers' prediction bias was greater in the failure

condition than in the success condition, revealing a unique effect of failed helping and indicating that social desirability cannot fully explain helpers' prediction bias in well-intentioned but counterproductive helping. Other psychological mechanisms unique to this context must exist.

3 Study 2: Parking a Car

Study 2 examined whether helpers' prediction bias in well-intentioned but counterproductive helping would emerge stably in a parking scenario and whether the form of helping would affect the bias. Helping can be proactive or reactive. Proactive helping occurs when helpers offer assistance before help recipients request it, whereas reactive helping occurs when helpers agree to help after receiving a request. These two forms may elicit different psychological responses—for example, help recipients may resist proactive help, perceiving it as infringing on their autonomy (Spitzmuller & Van Dyne, 2013). Study 2 examined whether helpers' prediction bias in well-intentioned but counterproductive helping exists across both helping forms.

3.1 Study 2a

3.1.1 Participants Five hundred twenty-nine participants (229 males, 300 females) with a mean age of 22.70 years (SD = 2.79) completed the study. The study employed a 2 (outcome: success/failure) × 2 (role: helper/help recipient) × 2 (form: proactive/reactive) between-subjects design.

3.1.2 Procedure Participants in the helper condition read a scenario imagining their friend Xiao Wang having difficulty parking. In the proactive condition, participants offered to help Xiao Wang park. In the reactive condition, Xiao Wang asked participants if they could help park, and participants agreed. In the success condition, participants were told they were very careful and the parking process went smoothly without any scratches. In the failure condition, participants were told that despite being careful, they scratched another car.

Participants in the help recipient condition read a similar scenario but imagined themselves as Xiao Wang receiving help from their friend. After reading the scenario, helpers predicted help recipients' reactions and help recipients rated their own reactions, including gratitude, satisfaction, willingness to seek help again, and recommendation willingness. The items were similar to those in Study 1, with "Xiao Zhang" replaced by "Xiao Wang." Finally, participants reported their gender and age.

3.1.3 Results and Discussion In Study 1, results were very similar across the four dependent measures. Therefore, given space constraints, we standardized each participant' s ratings across the four measures and averaged them to create a composite score, where higher scores indicated more positive predictions or evaluations. We report results for the composite score in the main

text and present means, standard deviations, and statistical tests for the four individual measures in Table 2 .

First, to examine whether helping form affected the composite score, we conducted a 2 (outcome) × 2 (role) × 2 (form) ANOVA. All interactions involving form were non-significant (ps > 0.190), indicating that helping form did not jointly influence the composite score with outcome and role. Therefore, we did not consider helping form in subsequent analyses.

Second, we conducted a 2 (outcome) × 2 (role) ANOVA on the composite score. The main effect of outcome was significant, $F(1, 525) = 1288.99$, $p < 0.001$, $p^2 = 0.71$. Participants predicted or rated higher scores in the success condition ($M = 0.78$, $SD = 0.41$) than in the failure condition ($M = -0.72$, $SD = 0.57$). The main effect of role was significant, $F(1, 525) = 44.53$, $p < 0.001$, $p^2 = 0.08$. Help recipients' actual scores ($M = 0.18$, $SD = 0.82$) were higher than helpers' predictions ($M = -0.16$, $SD = 0.94$).

Most importantly, the outcome × role interaction was significant, $F(1, 525) = 15.49$, $p < 0.001$, $p^2 = 0.03$. In the success condition, helpers underestimated help recipients' scores, $F(1, 252) = 4.83$, $p = 0.029$, $p^2 = 0.02$. In the failure condition, helpers underestimated help recipients' scores to a greater extent, $F(1, 273) = 47.42$, $p < 0.001$, $p^2 = 0.15$. This result indicates that prediction bias was greater in well-intentioned but counterproductive helping than in successful helping.

Study 2a replicated Study 1' s findings, showing that in well-intentioned but counterproductive helping, helpers overestimated help recipients' negative reactions (i.e., underestimated their positive reactions), providing further support for Hypothesis 1. In both success and failure conditions, helpers' prediction scores were lower than help recipients' actual ratings, ruling out the spotlight effect explanation. Additionally, on some measures, helpers underestimated help recipients' positive reactions even in the success condition, which may be due to help recipients' social desirability. However, helpers' prediction bias was greater in the failure condition than in the success condition, revealing a unique effect of failed helping and again indicating that social desirability cannot fully explain helpers' prediction bias in well-intentioned but counterproductive helping. Other causes must exist.

Notably, on the willingness to seek help again and recommendation willingness measures, both helpers' predictions and help recipients' ratings approached the scale ceiling. Therefore, we could not determine whether the near-accurate predictions in the success condition were due to ceiling effects or whether helpers could genuinely predict help recipients' reactions accurately. To address this issue, Study 2b used an unbounded measurement method to again compare helpers' prediction bias across success and failure conditions in the parking scenario.

3.2 Study 2b

3.2.1 Participants Two hundred eighty-one participants (149 males, 132 females) with a mean age of 27.58 years (SD = 5.20) completed the study. The study employed a 2 (outcome: success/failure) × 2 (role: helper/help recipient) between-subjects design.

3.2.2 Procedure In the first part of the study, participants in the helper condition imagined that their friend Xiao A returned after being out all day feeling very thirsty, and they offered Xiao A a bottle of water. Participants considered how this action would make Xiao A feel toward them and assigned this feeling the number "5" on a coordinate axis, where negative values represented negative feelings (lower values = more negative) and positive values represented positive feelings (higher values = more positive).

Participants in the help recipient condition imagined themselves as Xiao A returning thirsty after being out all day and receiving a bottle of water from friend Xiao A. They similarly assigned their feeling the number "5" on the same coordinate axis.

In the second part, participants in the helper condition read the parking scenario from Study 2a (reactive helping, outcome either success or failure). After reading, they used the water-offering event from Part 1 as a reference to predict help recipients' reactions. If participants thought the parking event would produce positive feelings, they provided a score greater than 0 (higher = more positive). For example, if the parking event produced positive feelings equal to the water-offering event, they would score it 5; if more positive, greater than 5; if less positive but still positive, between 0 and 5. If the parking event would produce negative feelings, they provided a score less than 0 (lower = more negative). If neutral, they scored it 0.

Participants in the help recipient condition read the same parking scenario and used the water-offering event as a reference to rate their own reactions. Thus, in Study 2b, participants' scores were unbounded, eliminating ceiling and floor effects. Finally, participants reported their gender and age.

3.2.3 Results and Discussion Following Jung et al. (2020), we winsorized extreme scores by replacing values below the 2.5th percentile with the 2.5th percentile value and values above the 97.5th percentile with the 97.5th percentile value.

We conducted a 2 (outcome) × 2 (role) ANOVA on participants' scores. The main effect of outcome was significant, $F(1, 277) = 25.82$, $p < 0.001$, $p^2 = 0.09$. Participants predicted or rated higher scores in the success condition ($M = 14.77$, $SD = 22.90$) than in the failure condition ($M = 3.20$, $SD = 13.35$). The main effect of role was not significant, $F(1, 277) = 0.76$, $p = 0.384$, $p^2 < 0.01$.

The outcome × role interaction was significant, $F(1, 277) = 4.85$, $p = 0.029$, $p^2 = 0.02$. In the success condition, helpers ($M = 16.24$, $SD = 24.58$) accurately predicted help recipients' scores ($M = 13.29$, $SD = 21.15$), $F(1, 138) = 0.58$, $p = 0.447$, $p^2 < 0.01$. In the failure condition, helpers ($M = 0.05$, $SD = 4.57$) underestimated help recipients' scores ($M = 6.89$, $SD = 18.43$), $F(1, 139) = 9.75$, $p = 0.002$, $p^2 = 0.07$. This result indicates that when ceiling and floor effects are eliminated, helpers' prediction bias exists only in well-intentioned but counterproductive helping, not in successful helping, revealing the uniqueness of this bias in failure contexts.

In summary, Studies 1 and 2 support Hypothesis 1, demonstrating that helpers exhibit prediction bias in well-intentioned but counterproductive helping—they underestimate help recipients' positive reactions. This bias occurs across both proactive and reactive helping, rules out the spotlight effect explanation, and indicates that social desirability cannot fully explain helpers' prediction bias in well-intentioned but counterproductive helping. Other more important causes must exist.

4 Study 3: Serving Guests

In the first two studies, we used gratitude, satisfaction, willingness to seek help again, and recommendation willingness as dependent measures. In Study 3, we used a more objective, interest-related measure—money—to further test helpers' prediction bias in well-intentioned but counterproductive helping. We used a guest-serving scenario to test the stability of prediction bias.

4.1 Participants

Two hundred sixty-four participants (135 males, 129 females) with a mean age of 28.44 years ($SD = 4.66$) completed the study. The study employed a 2 (outcome: success/failure) × 2 (role: helper/help recipient) between-subjects design.

4.2 Procedure

Participants in the helper condition read a scenario imagining themselves dining at a small restaurant that was very busy. The owner, overwhelmed with work, asked them to help serve porridge to several guests because the porridge barrel was near their seat. The participant agreed. In the success condition, participants were told they successfully served the porridge without spilling any. In the failure condition, participants were told they knocked over the porridge barrel, spilling it everywhere and breaking several bowls.

Participants in the help recipient condition read a similar scenario but imagined themselves as the restaurant owner receiving help from a guest. After reading the scenario, helpers in the success condition were told their meal cost 100 yuan and were asked to predict how much discount the owner would give them as a token of gratitude: "How much discount do you think the owner would

give you?" Help recipients rated how much discount they would actually give: "How much discount would you give this guest?" Helpers in the failure condition were told the loss amounted to 100 yuan and were asked to predict how much compensation the owner would claim: "How much compensation do you think the owner would ask from you?" Help recipients rated how much they would actually claim: "How much compensation would you ask from this guest?" Finally, participants reported their gender and age.

4.3 Results and Discussion

We conducted a 2 (outcome) × 2 (role) ANOVA on monetary amounts. The main effect of role was significant, $F(1, 260) = 9.62$, $p = 0.002$, $p^2 = 0.04$. Helpers' predictions ($M = 28.77$, $SD = 25.93$) were higher than help recipients' actual reactions ($M = 19.45$, $SD = 21.84$). The main effect of outcome was not significant, $F(1, 260) = 2.97$, $p = 0.086$, $p^2 = 0.01$.

The outcome × role interaction was significant, $F(1, 260) = 24.01$, $p < 0.001$, $p^2 = 0.08$. In the success condition, helpers ($M = 18.85$, $SD = 18.45$) accurately predicted the discount amount help recipients would give ($M = 23.93$, $SD = 20.64$), $F(1, 126) = 2.13$, $p = 0.147$, $p^2 = 0.02$. In the failure condition, helpers ($M = 37.54$, $SD = 28.45$) overestimated the compensation amount help recipients would claim ($M = 14.97$, $SD = 22.24$), $F(1, 134) = 26.46$, $p < 0.001$, $p^2 = 0.16$. This result indicates that prediction bias exists only in well-intentioned but counterproductive helping—helpers overestimated help recipients' compensation claims.

In summary, Study 3 replicated the results of Studies 1 and 2 using a monetary indicator, showing that prediction bias exists only in well-intentioned but counterproductive helping. In the next study, we explore the cause of this prediction bias.

5 Study 4: Moving Boxes

Study 4 used a box-moving scenario to test the stability of prediction bias and explore its underlying cause. As described above, compared with helpers' predictions, help recipients actually pay more attention to helpers' warmth and less attention to helpers' competence, leading helpers to overestimate help recipients' negative reactions. However, participants find it difficult to introspect about how much they focus on certain aspects. Following Kumar and Epley (2018), we measured participants' perceptions of helpers' warmth (reflected by helpfulness, Wang et al., 2017) and competence. Because well-intentioned but counterproductive helping inherently has low competence and high warmth attributes, lower competence ratings may reflect greater focus on competence, while higher warmth ratings may reflect greater focus on warmth.

5.1 Participants

Two hundred eighty-five participants (116 males, 169 females) with a mean age of 29.38 years (SD = 8.28) completed the study. The study employed a 2 (outcome: success/failure) × 2 (role: helper/help recipient) between-subjects design.

5.2 Procedure

Participants in the helper condition imagined themselves going home with their classmate Xiao Sun, who had placed fragile items in a suitcase. At the subway station, the elevator was under maintenance, and Xiao Sun needed help carrying the heavy suitcase upstairs. Worried about damaging the items, Xiao Sun asked the participant for help, and the participant agreed. In the success condition, participants were told the moving process went smoothly and the items remained intact. In the failure condition, participants were told there were some bumps during moving and the items in the suitcase broke.

Participants in the help recipient condition read a similar scenario but imagined themselves as Xiao Sun receiving help from their classmate. After reading the scenario, helpers first predicted help recipients' perceptions of their competence ("How competent do you think Xiao Sun would think you are," 1 = very incompetent, 7 = very competent) and warmth ("How helpful do you think Xiao Sun would think you are," 1 = not at all helpful, 7 = very helpful). Then helpers predicted help recipients' gratitude, satisfaction, willingness to seek help again, and recommendation willingness. Items were similar to those in Study 1, with "Xiao Zhang" replaced by "Xiao Sun."

Help recipients first rated helpers' competence ("How competent do you think Xiao Sun is," 1 = very incompetent, 7 = very competent) and warmth ("How helpful do you think Xiao Sun is," 1 = not at all helpful, 7 = very helpful). Then they rated their own gratitude, satisfaction, willingness to seek help again, and recommendation willingness. Items were similar to those in Study 1, with "Xiao Zhang" replaced by "Xiao Sun." Finally, participants reported their gender and age.

5.3.1 Dependent Variables As in Study 2a, we standardized each participant' s ratings across the four measures and averaged them to create a composite score, where higher scores indicated more positive predictions or evaluations. We report results for the composite score in the main text and present means, standard deviations, and statistical tests for the four individual measures in Table 3 .

A 2 (outcome) × 2 (role) ANOVA on the composite score revealed a significant main effect of outcome, $F(1, 281) = 323.10$, $p < 0.001$, $p^2 = 0.54$. Participants predicted or rated higher scores in the success condition (M = 0.63, SD = 0.40) than in the failure condition (M = -0.64, SD = 0.80). The main effect of role was significant, $F(1, 281) = 26.95$, $p < 0.001$, $p^2 = 0.09$. Help recipients' actual

scores (M = 0.18, SD = 0.81) were higher than helpers' predictions (M = -0.19, SD = 0.94).

Most importantly, the outcome × role interaction was significant, $F(1, 281) = 12.84$, $p < 0.001$, $p^2 = 0.04$. In the success condition, helpers accurately predicted help recipients' scores, $F(1, 141) = 2.90$, $p = 0.091$, $p^2 = 0.02$. In the failure condition, helpers underestimated help recipients' scores, $F(1, 140) = 24.64$, $p < 0.001$, $p^2 = 0.15$. This result indicates that helpers' prediction bias existed only in well-intentioned but counterproductive helping, supporting Hypothesis 1.

5.3.2 Mediating Variables A 2 (outcome) × 2 (role) ANOVA on competence ratings (Table 3) revealed a significant main effect of outcome, $F(1, 281) = 307.39$, $p < 0.001$, $p^2 = 0.52$. Participants predicted or rated higher competence in the success condition (M = 5.88, SD = 0.84) than in the failure condition (M = 3.74, SD = 1.25). The main effect of role was significant, $F(1, 281) = 16.30$, $p < 0.001$, $p^2 = 0.06$. Help recipients' ratings of helpers' competence (M = 5.06, SD = 1.43) were higher than helpers' predictions (M = 4.56, SD = 1.55).

The outcome × role interaction was significant, $F(1, 281) = 4.31$, $p = 0.039$, $p^2 = 0.02$. In the success condition, helpers accurately predicted help recipients' competence ratings, $F(1, 141) = 2.90$, $p = 0.086$, $p^2 = 0.02$. In the failure condition, helpers underestimated help recipients' competence ratings, $F(1, 140) = 13.74$, $p < 0.001$, $p^2 = 0.09$. This result indicates that in the failure condition, helpers underestimated help recipients' perceptions of their competence, suggesting that help recipients cared less about competence than helpers predicted.

A 2 (outcome) × 2 (role) ANOVA on warmth ratings (Table 3) revealed a significant main effect of outcome, $F(1, 281) = 53.55$, $p < 0.001$, $p^2 = 0.16$. Participants predicted or rated higher warmth in the success condition (M = 6.50, SD = 0.72) than in the failure condition (M = 5.68, SD = 1.18). The main effect of role was significant, $F(1, 281) = 17.21$, $p < 0.001$, $p^2 = 0.06$. Help recipients' ratings of helpers' warmth (M = 6.32, SD = 0.91) were higher than helpers' predictions (M = 5.86, SD = 1.15).

The outcome × role interaction was significant, $F(1, 281) = 9.47$, $p = 0.002$, $p^2 = 0.03$. In the success condition, helpers accurately predicted help recipients' warmth ratings, $F(1, 141) = 0.97$, $p = 0.325$, $p^2 < 0.01$. In the failure condition, helpers underestimated help recipients' warmth ratings, $F(1, 140) = 18.42$, $p < 0.001$, $p^2 = 0.12$. This result indicates that in the failure condition, helpers underestimated help recipients' perceptions of their warmth, suggesting that help recipients cared more about warmth than helpers predicted.

5.3.3 Mediation Effects We conducted a bootstrap mediation analysis with 5,000 samples, with role as the independent variable, competence as the media-

tor, outcome as the moderator, and composite score as the dependent variable. In the success condition, the 95% confidence interval was [-0.01, 0.11], containing zero, indicating that the mediation effect of competence was not significant. In the failure condition, the 95% confidence interval was [0.16, 0.52], not containing zero, indicating that the mediation effect of competence was significant (Figure 1 [Figure 1: see original paper]).

We conducted the same analysis with warmth as the mediator. In the success condition, the 95% confidence interval was [-0.04, 0.12], containing zero, indicating that the mediation effect of warmth was not significant. In the failure condition, the 95% confidence interval was [0.11, 0.37], not containing zero, indicating that the mediation effect of warmth was significant (Figure 2 [Figure 2: see original paper]).

These results show that first, helpers showed no prediction bias in the success condition but underestimated help recipients' evaluations in the failure condition, supporting Hypothesis 1. Second, helpers' prediction bias in the failure condition occurred because, compared with helpers' predictions, help recipients cared less about helpers' competence and more about helpers' warmth, supporting Hypothesis 2. In this study, we indirectly reflected participants' focus on competence and warmth through their perceptions. In Study 5, we will directly measure participants' focus on competence and warmth using a query method.

6 Study 5: Real-Time Thoughts

Study 4 preliminarily identified warmth and competence as causes of helpers' prediction bias. Study 5 records and analyzes participants' thoughts during prediction or judgment to further examine whether helpers and help recipients differ in their focus on helpers' competence and warmth.

According to query theory, people' s preferences are reflected in their thought processes (Johnson et al., 2007). Query theory has three assumptions. First, people break down decision problems into several queries. For example, when thinking about how grateful they are to a helper, help recipients might break the question into "What are the helper' s intentions?" and "How competent is the helper?" Second, people conduct queries in a particular order, with earlier queries reflecting more important considerations. If help recipients care about the helper's intentions, they will first consider "What are the helper's intentions?" and then "How competent is the helper?" If they care about competence, they will consider these in reverse order. Third, earlier queries generate more content. If help recipients first consider "What are the helper' s intentions?" they will generate more thoughts about intentions and fewer about competence.

Previous research has found that recording participants' thoughts during judgment and decision-making and calculating query order and query content indices can reflect the factors participants value (Lu & Xie, 2014). According to Hypothesis 2, help recipients should consider helpers' warmth earlier and more frequently than helpers predict, and consider helpers' competence later and less

frequently. Additionally, because previous studies showed that helpers' prediction bias is unique to well-intentioned but counterproductive helping compared with successful helping, this study only examines the failure condition.

6.1 Participants

One hundred twenty-three participants (49 males, 74 females) with a mean age of 21.36 years (SD = 2.67) completed the study. Participants were randomly assigned to either the helper or help recipient condition.

6.2 Procedure

Participants read the box-moving scenario from Study 4 and wrote down three to six thoughts that came to mind (Lu & Xie, 2014). Then helpers predicted help recipients' reactions and help recipients rated their own reactions, including gratitude, satisfaction, willingness to seek help again, and recommendation willingness. Items were identical to those in Study 4. Finally, participants reported their gender and age.

6.3 Coding

Two coders blind to the hypotheses categorized each thought as: related to helper' s competence (reflecting helping outcome), related to helper' s warmth (reflecting helping intention), or other. Examples are shown in Table 4 .

The coders agreed on the categorization of 86.60% of thoughts. Disagreements were resolved by a third coder. We then removed "other" thoughts and retained only competence- and warmth-related thoughts. For each participant, we calculated query order and query content indices using formulas (1) and (2) (Johnson et al., 2007; Lu & Xie, 2014).

Query Order Index = 2×(MR_{warmth} - MR_{competence}) / (n + 1)
Query Content Index = (n_{competence} - n_{warmth}) / n

Where MR_{warmth} is the median rank of warmth-related thoughts, MR_{competence} is the median rank of competence-related thoughts, n is the total number of thoughts listed, n_{competence} is the number of competence-related thoughts, and n_{warmth} is the number of warmth-related thoughts. Both indices range from -1 to 1. A query order index of 1 indicates all competence-related thoughts appeared before all warmth-related thoughts; -1 indicates the opposite. A query content index of 1 indicates all thoughts were competence-related; -1 indicates all were warmth-related. If a participant listed s thoughts related to only one dimension, we used a conservative approach by setting the median rank for the other dimension to s+1 and setting n to s+1.

Sixteen participants listed thoughts unrelated to helper's competence or warmth and were excluded from subsequent analyses of competence and warmth.

6.4 Results and Discussion

6.4.1 Dependent Variables As in Study 2a, we standardized each participant's ratings across the four measures and averaged them to create a composite score. We report results for the composite score in the main text and present means, standard deviations, and statistical tests for the four individual measures in Table 5 . Helpers underestimated help recipients' composite scores, $F(1, 121) = 11.37$, $p = 0.001$, $p^2 = 0.09$, demonstrating helpers' prediction bias.

6.4.2 Mediating Variables Helpers' query order index was higher than help recipients' index (Table 5), $F(1, 105) = 8.23$, $p = 0.005$, $p^2 = 0.07$, indicating that helpers thought about their competence earlier and warmth later than help recipients. Helpers' query content index was also higher than help recipients' index, $F(1, 105) = 8.24$, $p = 0.005$, $p^2 = 0.07$, indicating that helpers generated more thoughts about competence and fewer about warmth than help recipients. These results suggest that compared with help recipients' actual focus, helpers believed help recipients focused more on helpers' competence and less on helpers' warmth.

6.4.3 Mediation Effects We conducted bootstrap mediation analyses with 5,000 samples, with role as the independent variable, query order index as the mediator, and composite score as the dependent variable. The 95% confidence interval was [0.01, 0.24], not containing zero, indicating that query order index significantly mediated the effect (Figure 3 [Figure 3: see original paper]). We conducted the same analysis with query content index as the mediator. The 95% confidence interval was [0.01, 0.24], not containing zero, indicating that query content index also significantly mediated the effect (Figure 3).

These results indicate that compared with helpers' predictions, help recipients valued helpers' competence less and warmth more, and this difference caused helpers to underestimate help recipients' evaluations. Study 5 provides support for Hypothesis 2 by recording and analyzing helpers' and help recipients' real-time thoughts.

7 General Discussion

This research identified helpers' prediction bias about help recipients in well-intentioned but counterproductive helping and revealed its psychological cause. Studies 1, 2a, 2b, and 3 showed that when help failed, helpers overestimated help recipients' negative reactions (i.e., underestimated their positive reactions). This prediction bias emerged consistently across both proactive and reactive helping. The bias was unique: when help succeeded, helpers showed no prediction bias or only weak bias. Neither social desirability nor the spotlight effect could adequately explain this prediction bias. Studies 4 and 5 measured participants' focus on helpers' competence and warmth in different ways and found that when predicting help recipients' reactions, helpers believed help recipients cared about

helpers' competence, whereas when rating their own reactions, help recipients actually cared about helpers' warmth. This difference caused the prediction bias.

7.1 Theoretical Implications

Existing research on prediction bias in interpersonal interactions has focused primarily on non-conflict events, such as expressing gratitude (Kumar & Epley, 2018), conversing with others (Boothby et al., 2018), telling stories (Cooney et al., 2017), and gift-giving (Goodman & Lim, 2018; Kupor et al., 2017). These events have minimal potential to cause harm to others. In helping contexts, researchers have also focused on successful helping events (Zhang & Epley, 2009), which do not create interpersonal conflict. Does prediction bias exist in contexts that may cause interpersonal conflict? Does prediction bias in conflict contexts have unique manifestations compared with non-conflict contexts? Currently, only Levine and Cohen (2018) have examined this issue, finding that for honest communication—a potentially conflict-inducing event—people overestimate the negative impact on listeners. This occurs because communicators believe listeners care greatly about the specific content of the honesty while neglecting that listeners also care about the purpose of the communication.

Unlike Levine and Cohen (2018), this research focuses on counterproductive helping—a potentially conflict-inducing event—and examines helpers' and help recipients' focus on warmth and competence. The results show that helpers underestimate help recipients' positive reactions, and this prediction bias is unique: it disappears or weakens in the corresponding non-conflict event (successful helping). This result suggests that conflict and non-conflict contexts are asymmetrical, with conflict contexts being particularly unique. Future research could further examine prediction bias in conflict contexts and compare it with bias in non-conflict contexts to investigate whether they differ in magnitude or direction.

Furthermore, this research extends behavioral decision theory. Negativity bias—whereby people are more sensitive to negative than positive information (Baumeister et al., 2001)—is an important component of behavioral decision theory, with loss aversion being a specific manifestation (Kahneman & Tversky, 1979). This research reveals that people' s prediction processes are more sensitive to conflict events than non-conflict events, showing greater prediction bias in conflict contexts. This finding extends the applicability of negativity bias theory to prediction bias.

7.2 Practical Implications

The consequences of well-intentioned but counterproductive helping should not be underestimated. On one hand, it causes physical or material losses to help recipients, which are easily recognized. On the other hand, it may cause hidden psychological losses to helpers. For example, helpers may worry about receiving

negative evaluations from help recipients and become reluctant to help them again in the future, or even generalize this concern to other help-seekers. This "once bitten, twice shy" effect may contribute to a cold social atmosphere. These psychological losses and their consequences are often overlooked. Understanding the prediction bias revealed in this research and knowing help recipients' actual reactions may help reduce helpers' concerns about negative evaluations after counterproductive helping, alleviate their worries about relationship damage, and encourage them to continue helping enthusiastically in the future.

In addition to "knowing others," helpers must also "know themselves." This research finds that helpers overestimate others' negative reactions because they focus excessively on their own competence while neglecting their warmth. To correct this prediction bias, helpers need to value their warmth. Specifically, helpers can prioritize and think more about the warmth they demonstrate (Lu & Xie, 2014), thereby increasing their consideration of this dimension. For help recipients, even after being helped counterproductively, they should express gratitude for the helper' s good intentions, letting helpers realize that people care not only about competence but especially about intentions.

This research also has implications for children' s education. During children' s development, we need to protect their kind hearts, especially for more sensitive children. Helping children understand this prediction bias may reduce their negative experiences after counterproductive helping and encourage them to continue showing kindness to others.

7.3 Limitations and Future Directions

This research examined helping behaviors that are generally considered well-intentioned. In real life, some people may harm others under the guise of helping. Although this does not fall within our research scope, could help recipients in our studies have perceived the help as malicious? First, Study 4' s results show that help recipients' warmth ratings of helpers were significantly above the midpoint of 4 ($p < 0.001$), indicating that help recipients perceived helpers' intentions as good. Second, among the 259 queries listed by help recipients in Study 5, only two mentioned malicious intentions. This suggests that the vast majority of participants did not perceive the scenario as intentional harm. Future research could examine prediction bias in situations where people maliciously harm others under the guise of helping.

This research did not consider how the relationship closeness between interaction partners affects the magnitude of prediction bias. In our scenarios, the interaction partners were friends or classmates. When the counterproductive helper is a stranger, help recipients' reactions might differ. For strangers' help, help recipients might perceive greater warmth and thus show stronger positive reactions, while helpers might be especially concerned about their competence in front of strangers and thus incorrectly believe strangers also focus on their competence. In this case, the magnitude of prediction bias would be larger.

Additionally, this research did not involve situations where counterproductive helping has severe consequences. If consequences were severe, both helpers' and help recipients' evaluations might decrease, but the relative difference between them should remain stable. Because regardless of consequence severity, the difference in how much helpers and help recipients value competence versus warmth should persist. Future research could test whether our findings apply when counterproductive helping has severe consequences.

This research used hypothetical scenarios, asking participants to imagine experiencing them and then make predictions or judgments. We did not examine helpers' prediction bias in actual helping behaviors. We speculate that after actually causing counterproductive help, helpers would more strongly experience their own lack of competence and thus exaggerate help recipients' negative reactions to a greater extent. To increase ecological validity, future research could create real counterproductive helping situations.

This research relied on participants' self-reports. Future research could use more objective indicators, such as recording participants' physiological responses, to further test prediction bias in well-intentioned but counterproductive helping. Additionally, self-report measures may cause common method bias. However, this research used experimental designs with different participant groups, which helps control for common method bias.

Helping behavior is a form of altruism. Future research could examine whether similar prediction bias exists in other forms of altruistic behavior. Finally, people' s reluctance to help others may have many causes. This research provides one explanation from the perspective of prediction bias. Clearly, interventions based solely on this perspective are insufficient. Future research should continue exploring other ways to promote helping behavior.

7.4 Conclusion

In well-intentioned but counterproductive helping, helpers overestimate help recipients' negative reactions and underestimate their positive reactions. This prediction bias emerges consistently across various counterproductive helping contexts, whether helpers offer assistance proactively or respond to requests. Compared with successful helping, this prediction bias is unique—when help succeeds, helpers can accurately predict help recipients' reactions. The cause of this prediction bias is that helpers believe help recipients focus on helpers' competence, whereas help recipients actually focus on helpers' warmth.

References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, 93, 751–763.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5, 323–370.

Bohns, V. K., & Flynn, J. F. (2010). "Why didn't you just ask?"Underestimating the discomfort of help-seeking. *Journal of Experimental Social Psychology*, 46, 402–409.

Boothby, E. J., Cooney, G., Sandstrom, G. M., & Clark, M. S. (2018). The liking gap in conversations: Do people like us more than we think? *Psychological Science*, 29, 1742–1756.

Cooney, G., Gilbert, D. T., & Wilson, T. D. (2017). The novelty penalty: Why do people like talking about new experiences but hearing about old ones? *Psychological Science*, 28, 380–394.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878–902.

Flynn, F. J., & Bohns, V. K. (2008). If you need help, just ask: Underestimating compliance with direct requests for help. *Journal of Personality and Social Psychology*, 95, 128–143.

Garcia, S. M., Weaver, K., & Chen, P. (2019). The status signals paradox. *Social Psychological and Personality Science*, 10, 690–696.

Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one' s own action and appearance. *Journal of Personality and Social Psychology*, 78, 211–222.

Goodman, J. K., & Lim, S. (2018). When consumers prefer to give material gifts instead of experiences: The role of social distance. *Journal of Consumer Research*, 45, 365–382.

Hsee, C. K., & Tang, J. N. (2007). Sun and water: On a modulus-based measurement of happiness. *Emotion*, 7, 213–225.

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 461–474.

Jung, M. H., Moon, A., & Nelson, L. D. (2020). Overestimating the valuations and preference of others. *Journal of Experimental Psychology: General*, 149, 1193–1214.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Krueger, J., & Clement, R. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, 65, 596–610.

Kumar, A., & Epley, N. (2018). Undervaluing gratitude: Expressers misunderstand the consequences of showing appreciation. *Psychological Science*, 29, 1423–1435.

Kupor, D., Flynn, F., & Norton, M. I. (2017). Half a gift is not half-hearted: A giver–receiver asymmetry in the thoughtfulness of partial gifts. *Personality and Social Psychology Bulletin*, 43, 1–9.

Levine, E. E., & Cohen, T. R. (2018). You can handle the truth: Mispredicting the consequences of honest communication. *Journal of Experimental Psychology: General*, 147, 1400–1429.

Lu, J., & Xie, X. (2014). To change or not to change: A matter of decision maker' s role. *Organizational Behavior and Human Decision Processes*, 124, 47–55.

Newark, D. A., Bohns, V. K., & Flynn, F. J. (2017). A helping hand is hard at work: Help-seekers' underestimation of helpers' effort. *Organizational Behavior and Human Decision Processes*, 139, 18–29.

Reit, E. S., & Critcher, C. R. (2020). The commonness fallacy: Commonly chosen options have less choice appeal than people think. *Journal of Personality and Social Psychology*, 118, 1–21.

Scopelliti, I., Loewenstein, G., & Vosgerau, J. (2015). You call it "self-exuberance" ; I call it "bragging" : Miscalibrated predictions of emotional responses to self-promotion. *Psychological Science*, 26, 903–914.

Schroeder, J., Waytz, A., & Epley, N. (2017). Endorsing help for others that you oppose for yourself: Mind perception alters the perceived effectiveness of paternalism. *Journal of Experimental Psychology: General*, 146, 1106–1115.

Spitzmuller, M., & Van Dyne, L. (2013). Proactive and reactive helping: Contrasting the positive consequences of different forms of helping. *Journal of Organizational Behavior*, 34, 560–580.

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General*, 142, 151–162.

Wang, Y., & Xie, X. (2019). Prediction errors in helping and helping-seeking: Causes and coping. *Advances in Psychological Science*, 22, 571–579.

Wang, Z., Mao, H., Li, Y. J., & Liu, F. (2017). Smile big or not? Effects of smile intensity on perceptions of warmth and competence. *Journal of Consumer Research*, 43, 787–805.

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence and morality. *Journal of Personality and Social Psychology*,

67, 222–232.

Wojciszke, B. (2005). Morality and competence in person- and self-perception. *European Review of Social Psychology*, 16, 155–188.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *The Academy of Management Review*, 12, 250–264.

Zhang, Y., & Epley, N. (2009). Self-centered social exchange: Differential use of costs versus benefits in prosocial reciprocity. *Journal of Personality and Social Psychology*, 97, 796–810.

Appendices

A. Study 1 Materials: Teaching Roller Skating

[**Helper**] On Sunday, you and your friend Xiao Zhang go to the park together. The park is spacious and rents roller skates, so many people are skating. Xiao Zhang, who cannot skate, is interested but afraid of falling, so he asks you, who can skate, to teach him to prevent falls. You readily agree. While holding Xiao Zhang' s hand and teaching him to skate, you are very careful. Xiao Zhang skates smoothly without falling or getting injured. (However, while holding Xiao Zhang' s hand and teaching him to skate, you accidentally fall, causing Xiao Zhang to fall with you and suffer several scrapes.)

[**Help Recipient**] On Sunday, you and your friend Xiao Zhang go to the park together. The park is spacious and rents roller skates, so many people are skating. You, who cannot skate, are interested but afraid of falling, so you ask Xiao Zhang, who can skate, to teach you to prevent falls. Xiao Zhang readily agrees. While holding your hand and teaching you to skate, Xiao Zhang is very careful. You skate smoothly without falling or getting injured. (However, while holding your hand and teaching you to skate, Xiao Zhang accidentally falls, causing you to fall with him and suffer several scrapes.)

B. Study 2a Materials: Parking a Car

[**Helper-Proactive**] On Sunday, your friend Xiao Wang drives you to a restaurant for dinner. The restaurant is popular and it' s the weekend, so parking is scarce. After searching, Xiao Wang finds a spot between two cars. Seeing Xiao Wang' s difficulty parking, you offer to park for him. Because you are very careful, the parking goes smoothly without any scratches. (However, despite being careful, you scratch one of the cars.)

[**Helper-Reactive**] On Sunday, your friend Xiao Wang drives you to a restaurant for dinner. The restaurant is popular and it' s the weekend, so parking is scarce. After searching, Xiao Wang finds a spot between two cars. Xiao Wang complains about the difficulty and asks if you can park for him. You readily

agree. Because you are very careful, the parking goes smoothly without any scratches. (However, despite being careful, you scratch one of the cars.)

[**Help Recipient-Proactive**] On Sunday, you drive your friend Xiao Wang to a restaurant for dinner. The restaurant is popular and it' s the weekend, so parking is scarce. After searching, you find a spot between two cars. Seeing your difficulty parking, Xiao Wang offers to park for you. Because Xiao Wang is very careful, the parking goes smoothly without any scratches. (However, despite being careful, Xiao Wang scratches one of the cars.)

[**Help Recipient-Reactive**] On Sunday, you drive your friend Xiao Wang to a restaurant for dinner. The restaurant is popular and it' s the weekend, so parking is scarce. After searching, you find a spot between two cars. You complain about the difficulty and ask Xiao Wang if he can park for you. Xiao Wang readily agrees. Because Xiao Wang is very careful, the parking goes smoothly without any scratches. (However, despite being careful, Xiao Wang scratches one of the cars.)

C. Study 3 Materials: Serving Guests

[**Helper**] You frequently dine at a small restaurant. Today, while eating there, the restaurant is very busy and the owner is overwhelmed. Several guests ordered porridge. Because the porridge barrel is near your seat, the owner asks if you can help serve the guests. You readily agree and successfully serve the porridge without spilling any. (However, you knock over the porridge barrel, spilling it everywhere and breaking several bowls.)

[**Help Recipient**] A guest frequently dines at your small restaurant. Today, while eating there, your restaurant is very busy and you are overwhelmed. Several guests ordered porridge. Because the porridge barrel is near the guest' s seat, you ask if he can help serve the guests. The guest readily agrees and successfully serves the porridge without spilling any. (However, the guest knocks over the porridge barrel, spilling it everywhere and breaking several bowls.)

D. Study 4 Materials: Moving Boxes

[**Helper**] On vacation, you and your classmate Xiao Sun go home together. Xiao Sun placed some fragile items in a suitcase to take home. At the subway station, the elevator is under maintenance, and Xiao Sun needs to carry the heavy suitcase upstairs. Worried about damaging the items inside, Xiao Sun asks you to help carry the suitcase, and you readily agree. The moving process goes smoothly and the items remain intact. (However, there are some bumps during moving and the items in the suitcase break.)

[**Help Recipient**] On vacation, you and your classmate Xiao Sun go home together. You placed some fragile items in a suitcase to take home. At the subway station, the elevator is under maintenance, and you need to carry the heavy suitcase upstairs. Worried about damaging the items inside, you ask Xiao

Sun to help carry the suitcase, and he readily agrees. The moving process goes smoothly and the items remain intact. (However, there are some bumps during moving and the items in the suitcase break.)

Note: Figure translations are in progress. See original paper for figures.