Running Head: OVERCOMING THE OUTCOME BIAS

Overcoming the Outcome Bias: Making Intentions Matter

Ovul Sezer[1]

Ting Zhang[2]

Francesca Gino[1]

Michael I. Norton[1]

[1]*Harvard Business School*

Harvard University

[2]*Columbia Business School*

Columbia University

Abstract

People often make the well-documented mistake of paying too much attention to the outcomes of others' actions while neglecting information about the original intentions leading to those outcomes. In five experiments, we examine interventions aimed at reducing this outcome bias in situations where intentions and outcomes are misaligned. Participants evaluated an individual with fair intentions leading to unfavorable outcomes, an individual with selfish intentions leading to favorable outcomes, or both individuals jointly. Contrary to our initial predictions, participants weighed others' outcomes more—not less—when these individuals were evaluated jointly rather than separately (Experiment 1). Consequently, separate evaluators were more intention-oriented than joint evaluators when rewarding and punishing others (Experiment 2a) and assessing the value of repeated interactions with these individuals in the future (Experiment 2b). Third-party recommenders were less outcome-biased in allocating funds to investment managers when making separate evaluations relative to joint evaluations (Experiment 3). Finally, raising the salience of intentions prior to discovering outcomes helped joint evaluators overcome the outcome bias, suggesting that joint evaluation made attending to information about intentions more difficult (Experiment 4). Our findings bridge decision-making research on the outcome bias and management research on organizational justice by investigating the role of intentions in evaluations.

Keywords*:* outcome bias; intentions; joint evaluation; separate evaluation; judgment

Overcoming the Outcome Bias: Making Intentions Matter

Half of the results of a good intention are evil; half the results of an evil intention are good.

—Mark Twain, "The Dervish and the Offensive Stranger"

Consider a well-intentioned physician who conducts a thorough physical exam on a patient and discovers the patient has as a condition that could be potentially serious if left untreated. The physician prescribes medication recommended by the American Medical Association, but the patient later dies from extremely rare complications after taking the medication. Now imagine instead that a more selfish physician conducts a cursory physical on the same patient in order to leave work early and consequently overlooks the patient's condition. The condition remains untreated, and the patient lives for another 40 years without complications. As Twain's quote and these anecdotes suggest, "good" intentions do not necessarily lead to desired outcomes, and "bad" intentions do not necessarily lead to unfavorable outcomes. The well-intentioned physician could face a malpractice lawsuit, while the more selfish physician would not face any punishment. Our legal system reflects individuals' tendency to focus on outcomes when judging behavior (Baron & Hershey, 1988), but as prior research has shown, distributing rewards and punishment based on the desirability of end results can perpetuate suboptimal decision-making (Gino, Moore, & Bazerman, 2012; Levitt & Dwyer, 2002; Moore, Tetlock, Tanlu, & Bazerman, 2006).

Organizations can benefit from a better understanding of how to help individuals look beyond end results. The literature on the outcome bias concludes that individuals overweight outcome favorability when making evaluations (Allison, Mackie, & Messick, 1996; Baron &

Hershey, 1988; Hastie & Dawes, 2001). That is, people perceive the same decision (e.g., a surgeon decides to operate on a patient) to be lower in quality when it leads to a bad outcome (the patient dies) rather than a good outcome (the patient survives), all else being equal. In organizational contexts, employees are more likely to be satisfied with unfair procedures if their outcome is favorable to them than if it is unfavorable to them (Brockner & Wiesenfeld, 1996; Messick & Sentis, 1979). Thus, much of the research on the outcome bias has focused on how individuals neglect information about the quality of others' decisions (e.g., the surgeon's decision to operate) or the process that led to those outcomes (e.g., the fairness of procedures).

Whereas prior research in organizational justice has focused on how outcome favorability both influences and is influenced by perceptions of decision quality and procedural fairness (Brockner & Wiesenfeld, 1996; Leventhal, 1980; Lind & Tyler, 1988; Thibaut & Walker, 1975), we consider another important antecedent of outcomes: *intentions*. Judgments of others' intentionality—that is, whether their behavior is perceived to be goal-directed or accidental—influence people's perceptions (Dennett, 1987; Heider, 1958; Shultz & Wells, 1985) and determine the course of social interactions (Heider, 1958, Fiske, 1989, Shaver, 1985). We study situations in which individuals neglect information about a decision-maker's intentions, which can be crucial and relevant to evaluation when outcomes are idiosyncratic. In many cases, intentions and outcomes of decisions match: the outcome achieved fully reflects the individual's original intentions. However, in our fundamentally noisy and complex world, intentions and outcomes often diverge. In the current research, we examine such situations and investigate factors that may help individuals pay attention to others' intentions instead of the outcome of their decisions.

One possible strategy for reducing the outcome bias as a consequence of intention neglect entails evaluating different outcomes and the intentions that led to them jointly rather than separately. For instance, managers may choose to evaluate the performance reviews of multiple employees simultaneously rather than sequentially. Joint evaluation describes situations in which two different options are evaluated simultaneously, whereas separate evaluation describes contexts in which each option is presented and evaluated on its own (Bazerman, Loewenstein, & White, 1992; Hsee, Loewenstein, Blount, & Bazerman, 1999). Relative to separate evaluation, joint evaluation has been shown to increase reason-based decisions in both amoral (Bazerman, Moore, Tenbrunsel, Wade-Benzoni, & Blount, 1999) and moral contexts (Bazerman et al., 2011; Gino et al., 2012; Gino, Shu, & Bazerman, 2010). These findings suggest that joint evaluation may reduce the outcome bias relative to separate evaluation, especially in situations where outcomes and intentions are incongruent. Therefore, prior research would predict that joint evaluation of both bad intentions leading to good outcomes and good intentions leading to bad outcomes may mitigate the outcome bias relative to separate evaluation of each decision.

However, results from a pilot study we conducted revealed that when information about the intentions of the decision maker was varied, joint evaluation *exacerbated* rather than reduced the outcome bias in those who could reward or punish others. In this pilot study (replicated in Experiment 2a), participants evaluated (1) either a partner who made a fair decision that had the expected value of benefitting both the participant and the partner equally, (2) a partner who made a selfish decision that would likely benefit the partner more than the participant, or (3) both the fair and selfish partners jointly. Participants then learned that despite the fair partner's intentions, the participant received a worse outcome relative to the partner; in contrast, in spite of the selfish partner's unfair decision, the outcome favored the participant more than the partner (see

Appendix A in the Online Supplement for more details). In this situation where intentions and outcomes were misaligned, separate evaluators factored in their partners' intentions more when rewarding and punishing their partners relative to joint evaluators.

Based on the results of this pilot study, this paper presents five experiments showing the conditions in which joint evaluation enhances the outcome bias rather than reducing or eliminating it. In particular, we tested situations in which both the intention of the decision maker and the outcome of the decision differed across individuals. We show that joint evaluations repeatedly led to greater bias toward outcomes, whereas separate evaluations led participants to consider intentions more heavily. We then consider how joint evaluation enhances the outcome bias, extend these findings to managerial contexts where individuals evaluate employees with varied intentions and outcomes, and explore interventions aimed at mitigating the outcome bias in joint-evaluation contexts. We conclude by discussing the theoretical and practical implications of our findings.

Outcome Bias due to Neglect of Decision Quality and Procedures

From a rational perspective, individuals with full information about both the decisions and outcomes involved in a situation should not base their evaluation solely on outcomes, particularly in situations where outcomes reflect noise in the environment and provide no additional information about the quality of a decision (Bazerman & Moore, 2013; Hastie & Dawes, 2001). However, a long stream of research has shown that people consistently overweigh outcome information in their evaluations of decision quality. In a classic example, participants rated the quality of a surgeon's decision to perform a risky operation (Baron & Hershey, 1988). Although participants read about identical decision processes, they received different information about the outcome of the surgeon's decision. Participants who read that the patient

died soon after the surgery rated the surgeon's decision to operate to be of lower quality than did participants who read that the patient survived.

Research on procedural fairness has mainly focused on the *structural* processes (e.g., guidelines for selecting the decision maker, rules for ensuring that a decision maker does not abuse power, or procedures that allow change in allocations) used to arrive at a particular outcome (Brockner & Wiesenfeld, 1996; Lind & Tyler, 1988), leaving a gap in research on the interventions that help individuals more carefully consider others' *thought* processes, including the *intentions* behind their decisions. We seek to fill that gap with the current research.

Outcome Bias due to Intention Neglect

While prior research has examined the outcome bias as a result of neglecting decision quality and procedural fairness, this research focuses on the factors that influence whether individuals make outcome-biased evaluations as a result of neglecting important information about others' intentions. For example, managers frequently evaluate multiple employees, all of whom have different intentions that produce varied outcomes. When intentions and outcomes differ, neglect of intentions could lead people to reward individuals with bad intentions and punish those with good intentions.

Existing research suggests that individuals tend to ignore information about intentions when information about outcomes is available (Cushman, 2008; Cushman et al., 2009; Pizarro, Uhlmann & Bloom, 2003; Weiner, 1995). People are more likely to give themselves credit for their own good intentions—even if they fail to follow through on them (Kruger & Gilovich, 2004). However, when evaluating others, they are more focused on the outcomes reached than on the intentions behind those actions (Jones & Nisbett, 1971). In particular, when both intentions and outcomes differ, individuals tend to neglect information about intentions in the

presence of idiosyncratic outcomes (Cushman, 2008; Cushman et al., 2009; Pizarro, Uhlmann & Bloom, 2003; Weiner, 1995). Previous studies (e.g., Cushman et al., 2009; Stanca, 2010) that compared responses to identical outcomes reached by different combinations of intentional choices and chance found that accidental outcomes were the main factors guiding punishment.

Furthermore, related research in behavioral ethics and moral psychology has shown that outcome information prevents individuals from considering the actor's intentions to act ethically or unethically. For example, outcome information influences attributions of blame and responsibility (Alicke & Davis, 1989; Lowe & Medway, 1976; Schroeder & Linder, 1976) and ethical judgments; people evaluate actions based on ethically questionable intentions that lead to negative consequences as more unethical and blameworthy than the same actions when they generate positive outcomes (Gino et al., 2012; Gino et al., 2010).

Despite evidence that individuals neglect intentions when outcomes are present, little work has investigated interventions aimed at helping individuals consider others' intentions. Charness and Levine (2007) suggest that explaining what the intended outcomes might have been could mitigate intentions neglect. In their study, employees worked harder when their firm released information that it had good intentions to pay them well, even if actual pay was later reduced due to the firm's bad luck. Building on this research, we present five studies aimed at taking a broader approach to understanding the impact of intentions on evaluations of others' decisions that led to desirable or undesirable outcomes and investigate the conditions in which individuals tend to ignore intentions.

Joint Evaluation

Prior research on the outcome bias, distributive justice, and procedural fairness has primarily studied individuals' evaluations of *one* specific situation (Brockner & Wiesenfeld,

1996; Gilliland, 1994; Lind & Lissak, 1985, McFarlin & Sweeney, 1992). These studies compare individuals' reactions to an unfavorable outcome that arose from an unfair process against other respondents' reactions to an unfavorable outcome that arose from fair processes. That is, they investigate *separate evaluation*. However, managers and employees often evaluate *multiple* situations simultaneously—that is, they often engage in *joint evaluation*. For example, managers may decide whether to promote multiple employees who all have differing motivations for wanting the job and differing performance in their current roles.

These modes of evaluation—separate and joint—differentially impact the extent to which individuals make biased decisions. Individuals exhibit more rational preferences when evaluating options jointly than when evaluating options separately (Bazerman, Loewenstein, & White, 1992; Bazerman, Schroth, Shah, Diekmann, & Tenbrunsel, 1994; Bazerman, Tenbrunsel, & Wade-Benzoni, 1998; Zikmund-Fisher, Fagerlin, & Ubel, 2004). As a result, joint decision-making decreases gender-based discrimination in hiring decisions (Bohnet, Van Geen, & Bazerman, in press), promotes more optimal consumption choices (Hsee et al., 1999), and reduces the impact of victim identifiabilty on biased moral judgments (Gino et al., 2010).

One explanation offered by past research for findings of preference reversals under joint evaluation is that making decisions simultaneously is more likely to generate reflective and reason-based thinking, whereas making separate evaluations is more likely to generate choices based on affective responses (Bazerman, Tenbrunsel & Wade-Benzoni, 1998; Bazerman & Moore, 2013; Bazerman, et al., 2011). That is, individuals are more likely to make decisions based on their first impulses when presented with options separately rather than jointly (Milkman, Chugh, & Bazerman, 2009; Moore & Loewenstein, 2004). Another explanation is that by allowing individuals to consider various options simultaneously, joint evaluation provides

them with more information about the decision (Gino et al., 2010; Hsee et al., 1999), while separate evaluation only invokes internal reference points (Kahneman & Miller, 1986).

Prior work finding that joint evaluation debiased decisions focused on decisions in fairly simple contexts (Bazerman, Loewenstein & White, 1992; Nowlis & Simonson, 1997; Kahneman et al., 1993; Hsee et al., 1999; Gino et al., 2010). For instance, in original demonstration of preference reversals under joint evaluation, participants evaluated two possible negotiation outcomes—an equal split or a money-maximizing option—either separately or jointly. Similarly, in a study by Gino, Shu & Bazerman, (2010), participants evaluated the ethicality of different acts, holding the decision-making process and intentions of the actors constant. By making different alternatives available, joint evaluation provides a comparison point for evaluation and thus highlights keys differences on attributes that are easily overlooked in separate evaluation contexts.

However, in more complex contexts, when multiple attributes vary simultaneously, joint evaluation might lose that advantage. Although past research offers a compelling argument that joint evaluation is an effective tool for reducing biases, the *evaluability hypothesis* suggests that joint evaluation is not necessarily beneficial if simultaneous comparisons do not offer any additional information about the attributes of options or the relative importance of these attributes beyond what individuals can deduce from separate evaluation (Bazerman et al., 1999; Handgraaf, Van Dijk, Wilke, & Vermunt, 2004; Hsee, 1996; Wang & Murninghan, 2013). In contexts where joint evaluation provides information on multiple factors, it may actually increase bias. When intentions and outcomes vary, joint evaluation may divert attention away from intentions, which are more often difficult to infer, toward outcomes, which are more salient.

Our pilot study (replicated in Experiment 2a) provided initial evidence that joint evaluators are more outcome-biased relative to separate evaluators in contexts where intentions and outcomes are misaligned. We found that when making judgments of an individual with fair intentions leading to unfavorable outcomes or an individual with selfish intentions leading to favorable outcomes, separate evaluators judging only one of the two individuals were more likely to reward based on intentions, while joint evaluators judging both individuals were more likely to reward based on outcomes. In support of these findings from our pilot, studies in which individuals evaluated varying intentions and outcomes simultaneously under a strategy method (similar to joint-evaluation contexts) found individuals were more biased toward outcomes (Cushman et al., 2009), whereas a different study in which individuals evaluated varying intentions and outcomes on a case-by-case basis using the direct response method (similar to separate valuation) found greater reliance on intentions (Gerstenberg, Lagnado, & Kareev, 2010). Consistent with results from our pilot and previous studies, we test the following hypothesis: compared to separate evaluators, joint evaluators are less (more) likely to evaluate negatively (positively) and punish (reward) individuals who reach good outcomes despite selfish intentions.

Joint evaluation may prevent individuals from being able to recognize intentionality as the more relevant attribute when evaluating the decision maker. Thus, in situations where intentions and outcomes are misaligned, we predict that joint evaluators are more likely to ignore less salient but important information on intentions and instead focus on easier-to-evaluate factors that have more salient and clear consequences: outcomes.

Overview of Present Research

We conducted five experiments to test how two different modes of evaluation, joint and separate, influence the outcome bias. In Experiment 1, we tested whether joint evaluation would exacerbate the outcome bias in situations where intentions and outcomes are unconnected. In Experiment 2a, we explored how individuals respond to accidental outcomes under different modes of evaluation. In Experiment 2b, we tested whether the findings of Experiment 2a would hold when evaluators have to select a partner for additional games with concrete rewards based on their evaluation. In Experiment 3, we investigated whether these findings would hold when individuals either jointly or separately evaluated investment managers, demonstrating the organizational implications of our findings. Finally, in Experiment 4, we extended our investigation by requiring individuals to make preliminary reward decisions before outcomes were revealed and found that this intervention reduced the outcome bias in joint contexts.

Experiment 1: Outcome Bias under Joint vs. Separate Evaluation

Experiment 1 tested whether overreliance on outcomes in evaluation of others is more prevalent in joint evaluation relative to separate evaluation contexts. Based on findings in our pilot study, we predicted that in situations where intentions and outcomes are not aligned, the outcome bias would be stronger in joint evaluation than in separate evaluation contexts.

Method

Participants. One hundred and ninety-four individuals (51% male, 49% female, M_{age} = 28.45, *SD* = 10.59) from a pool of participants managed by a university in the northeastern United States participated in the study for $15.

Design and Procedure. Participants were randomly assigned to one of the three experimental conditions. Sixty-six participants evaluated a physician who had good intentions

that led to unfavorable patient outcomes, 63 participants evaluated a physician who had selfish intentions that led to favorable patient outcomes, and 65 participants evaluated these two physicians jointly. Participants were told the physicians could choose between two equally effective drugs to treat their patient (drug names were counterbalanced throughout the study):

> Both drugs [Xenosilycate and Zencomayacil] are equally effective in clinical tests. In the majority of the cases (95% of the time), patients recover fully within 24 hours. In very rare instances (5% of the time), both drugs can cause side effects, such as nausea, bleeding in the stomach, or loss of appetite. Extensive research has shown that there is absolutely no method to determine who will be influenced by side effects, so there are no tests that doctors can run to determine whether patients will experience these side effects ahead of time.

Although both drugs were equally effective based on clinical tests, one was cheaper for the patient, while the other would generate more revenue for the physician:

> The doctors receive different Medicare reimbursement fees for administering these drugs. The reimbursement for [Zencomayacil] is 20% higher than [Xenosilycate], so the physician pockets more money by prescribing [Zencomayacil]. However, this drug is more expensive for the patient.

Participants evaluating the well-intentioned physician learned that the doctor prescribed the cheaper drug to save the patient money. Despite these good intentions, the patient suffered from adverse side effects and spent the night at the hospital. Participants evaluating the physician with selfish intentions learned that the doctor prescribed the more expensive drug to generate more revenue for the clinic. Although the physician had selfish intentions, the patient made a full recovery without side effects or hospitalization. In the joint evaluation condition, participants evaluated both physicians. The order of the scenarios was counterbalanced.

Dependent measure. After learning the outcome of the drug treatment, participants imagined they had been given the opportunity to review the doctor on a crowd-sourced website tracking the quality of physicians (1 star = *very bad*, 5 stars = *very good*).

Results

Across all studies unless otherwise noted, we conducted linear mixed model analyses for continuous dependent variables and logistic mixed model analyses for binary variables, treating participants as random intercepts and the framing of evaluation (joint vs. separate), intention/outcome combinations (fair intentions leading to unfavorable outcomes vs. selfish intentions leading to favorable outcomes), and their interactions as fixed effects.[1] We unpack interactions with simple effects and present only relevant main effects. (For more details about all analyses conducted, see Appendix C in the Online Supplement.)

Physician ratings. We found a significant interaction between framing of evaluation and intentions, confirming that the difference in ratings for selfish and well-intentioned physicians differed more in the joint condition than in the separate evaluation condition, $F(1, 255) = 13.98$, $p < .001$, $\eta_p^2 = .05$. When participants evaluated physicians jointly, participants rated the selfish physician ($M = 4.14$, $SD = .89$) more positively than they did the well-intentioned physician ($M = 2.79$, $SD = 1.27$), $B = 1.35$, $t(255) = 7.03$, $p < .001$. However, when participants made their evaluations separately, ratings for the well-intentioned physician ($M = 3.33$, $SD = 1.18$) did not differ from those of the selfish physician ($M = 3.67$, $SD = 1.00$), $B = -.33$, $t(255) = 1.72$, $p = .09$, suggesting that separate evaluators were less outcome-biased than joint evaluators.

Insert Table 1 and Figure 1 about here

Discussion

Experiment 1 provided initial evidence that joint evaluators were more likely to neglect intentions and overweight outcomes as compared to separate evaluators. We found that joint evaluators evaluated good outcomes more favorably despite selfish intentions. In Experiments 2a

and 2b, we investigated how these findings extend beyond mere evaluations of individuals with selfish or good intentions and affect behaviors towards these individuals.

Experiment 2a: Rewarding and Punishing based on Accidental Outcomes

Experiment 1 tested how the mode of evaluation—joint or separate—influenced the extent to which individuals incorporated information about others' intentions in their evaluations. How do these evaluations then translate into behavior toward these individuals? To examine behavior toward these targets of evaluation, we adopted a paradigm developed by Cushman and colleagues (2009) that dissociates intentions from outcomes and allows individuals to reward and punish others based on information about others' intentions and outcomes. Evaluators can choose to reward or punish their partner based on the partner's decision to roll either a fair die or a selfish die that would determine the final payoffs for both players.

Because we are interested in the effect of evaluation mode on the outcome bias, we focus primarily on the difference in evaluations of partners that reached either 1) unfavorable outcomes for the evaluator despite selecting the fair die or 2) favorable outcomes for the evaluator despite selecting the selfish die. Rewarding favorable outcomes in spite of selfish intentions more than unfavorable outcomes despite fair intentions indicates an outcome bias, and comparing these differences in rewards between joint and separate modes of evaluation reveals the extent to which joint evaluators are comparatively outcome-biased relative to separate evaluators. We expected to replicate results from Experiment 1 showing that individuals are more outcome-biased under joint evaluation as compared to separate evaluation.

Method

Participants. Three hundred individuals (51% male, 49% female, M_{age} = 33, *SD* = 11.66) recruited through Amazon's Mechanical Turk participated in a 20-minute online study for $0.75, a standard rate for this market.

Design and procedure. We randomly assigned participants to one of three conditions: 99 participants evaluated a hypothetical partner with fair intentions that led to an unfavorable outcome for the participant, 100 participants evaluated an individual with selfish intentions that led to a favorable outcome for the participant, and 101 participants evaluated both partners with fair and selfish intentions leading to unfavorable and favorable outcomes jointly.

In all conditions, participants played the "die game," which generates accidentally generous or accidentally selfish outcomes (Cushman et al., 2009). Participants were asked to imagine that they had been paired with a partner whose task was to choose between two different dice—Die A (a "selfish" die) and Die B (a "fair" die)—and roll the selected die to determine the payoffs for the die selector and the participant. All participants were assigned to play the role of an evaluator (Player 2). In separate-evaluation conditions, participants imagined that they had one partner (Player 1), whereas in the joint-evaluation condition, participants imagined they had two partners (Player 1 and Player 3).

The choice between Die A and Die B signaled to participants whether their partner(s) had selfish or fair intentions (see Table 2 for payoff chart). Partners who chose Die A signaled selfish intentions, as the distribution of the outcomes was less favorable toward the participant: rolling 1, 2, 3, or 4 benefitted the die selector (partner) with $10, but left the participant with $0; rolling a 5 led to a fair allocation, leaving both the die selector and the participant with $5; and only a die roll of 6 benefitted the participant with $10, giving the die selector $0. Thus, selection of Die A signaled the partner's selfish motivations, as Die A had a higher probability of arriving at a

favorable allocation for the partner at the cost of reaching an unfavorable allocation for the participant. In contrast, Die B signaled fair intentions, as the distribution of outcomes equally favored both the die selector and the participant: rolling 1, 2, 3, or 4 led to fair outcomes, with both the die selector and the participant receiving $5; rolling a 5 favored the die selector with $10, while providing the participant with $0; and rolling a 6 favored the participant with $10, while giving the die selector $0.

Insert Table 2 about here

Those in the separate-evaluation condition were informed that their partner (Player 1) chose either Die A (the "selfish" die) or Die B (the "fair" die). Participants who were paired with partners who selected the "selfish" die imagined that the die landed on a 6, giving participants $10 and leaving the die selector with $0. In contrast, those paired with partners who selected the "fair" die imagined that the die landed on a 5, providing the partner with a generous $10 outcome and leaving the participant with $0. Thus, half the participants in the separate-evaluation condition were paired with partners whose actions had favorable outcomes for the participants despite selfish intentions, and the other half received undesirable outcomes despite their partner's fair intentions.

In the joint-evaluation condition, participants were paired with two partners who independently determined which die to roll. Participants were informed that one of their partners (Player 1) chose Die B (the "fair" die), which landed on a 5, leaving participants with $0 and this partner with $10; in contrast, they were informed that their other partner (Player 3) chose Die A (the "selfish" die), which landed on a 6, leaving this partner with nothing and giving participants $10. Thus, participants in the joint-evaluation condition made reward and punishment decisions

for both Player 1, who rolled an unfavorable outcome despite his fair intentions, and Player 3, who rolled a desirable outcome for the participant despite his selfish intentions. Participants first learned about both players' die selection on one screen. On the following screen, participants read information about the outcome of the selected die and then evaluated both players.

Dependent measures. After learning the outcome of the die selections, participants imagined they had been given the opportunity to reward or punish each partner with a maximum of $9. They were told the allocation would change only their partner's allocation and not their own. Our main dependent variables include 1) whether evaluators decided to reward or punish and 2) the allocation amount. We coded rewards as positive allocations and punishments as negative allocations. Those who evaluated both partners (Player 1 and Player 3) jointly could reward or punish both players up to $9 each. All participants were informed of this opportunity at the beginning of the study when they learned the rules of the game, and they were reminded about this opportunity after learning the outcome of the die roll(s).

Results

Allocations. We found a significant interaction between framing of evaluation and intentions, confirming that allocation of rewards for favorable and unfavorable outcomes differed in the joint condition but not in the separate condition, $F(1, 397) = 9.50$, $p = .002$, $\eta_p^2 = .02$.

When participants made decisions jointly, rewards differed based on outcomes, $B = -3.82$, $t(397) = 4.81$, $p < .001$. Participants rewarded the selfish partner (M = $4.20, $SD = 4.54$) more than they did the fair partner (M = $0.39, $SD = 5.64$). However, under separate evaluation, the amount of reward based on evaluation of the fair partner (M = $2.09, $SD = 6.43$) did not differ significantly from that of the selfish partner (M = $2.44, $SD = 5.82$), $B = -.34$, $t(397) = .43$, $p = .67$, suggesting that separate evaluators were less outcome-biased than joint evaluators.

Insert Table 3 and Figure 2 about here

Percent rewarding. A multi-level logistic regression revealed an interaction between the framing of evaluations and intentions on reward decisions, $B = 1.10$, Wald $\chi^2 = 2.58$, $p < .001$, odds ratio = 3.01. Joint evaluators were more likely to reward selfish individuals (72%) than fair individuals, (41%), $B = 1.37$, Wald $\chi^2 = 4.31$, $p < .001$, odds ratio = 3.94. Separate evaluators also rewarded selfish (57%) and fair partners (51%) in similar proportions, $B = .26$, Wald $\chi^2 = .92$, $p = .36$, odds ratio = 1.31.

Percent punishing. We also found an interaction between the framing of evaluations and intentions on punishment decisions, $B = -1.21$, Wald $\chi^2 = -2.38$, $p = .02$, odds ratio = .23. Joint evaluators were more likely to punish partners with fair intentions (33%) than those with selfish intentions (10%), $B = -1.48$, Wald $\chi^2 = -3.76$, $p < .001$, odds ratio = 1.66, whereas separate evaluators punished fair (29%) and selfish partners (24%) in similar proportions, $B = -.27$, Wald $\chi^2 = -.84$, $p = .40$, odds ratio = .76 (see Figure 2).

Discussion

Experiment 2a showed that joint evaluators were more likely to neglect intentions and overweight outcomes as compared to separate evaluators. More specifically, our results showed that joint evaluation of fair intentions and selfish intentions exacerbated the extent to which individuals rewarded and punished others' behaviors based on outcomes. Put differently, separate evaluators appeared less outcome-biased than joint evaluators. Relative to joint evaluators, separate evaluators were more likely to reward good intentions, even when the outcome was unfavorable.

The presence of the outcome bias in joint-evaluation contexts is consistent with results that Cushman et al. (2009) found under the strategy method. In their studies, all evaluators in a

within-subjects design essentially made joint evaluations as they rewarded or punished their partner across a variety of scenarios in which their partners' intentions and outcomes were either aligned or mismatched.

One possible limitation of Experiments 1 and 2a is that participants made their evaluations based on hypothetical scenarios. Experiment 2b investigates whether the results are generalizable to situations in which individuals' decisions have real consequences for their partners. Additionally, an alternative explanation of these findings is that participants were motivated to reward or punish their partners not for their intentions or outcomes, but instead to reduce inequity between workers. Because the purpose of rewarding or punishing partners was left to participants' own interpretations, joint evaluators may have been more motivated than separate evaluators to redistribute earnings between partners such that both partners would receive relatively equal pay. We address this possibility in Experiment 2b.

Experiment 2b: Outcome Bias in Repeated Interactions

In Experiment 2a, rewarding partners based on intentions was not necessarily the normative choice, as we did not provide a clear goal to participants about the purpose of the reward-and-punishment task. As a result, participants could have interpreted the reward-and-punishment task as a means to redistribute wealth. Thus, joint evaluators in Experiment 2a may have appeared more outcome-biased than separate evaluators because they were more sensitive to the inequity between their partners.

To address this confound, we developed a paradigm in which making decisions based on their partners' intentions was the more rational decision. After participants played the die game, as in Experiment 2a, they made decisions about whether they wanted to continue interacting with their partner in the future—a decision that should not be influenced by information about

idiosyncratic outcomes achieved in the prior game. In this new task, participants in Experiment 2b had the option to participate in a multi-round dictator game in which the same partner was the dictator across all rounds. Whereas the partners' intentions and outcomes were not necessarily aligned in the die game, they were perfectly aligned in this dictator game because the outcomes for the participant were exactly what their partner—the dictator—had intended. When evaluating whether to play this multi-round dictator game with their partner, participants benefited from ignoring information about outcomes from the die game in favor of information about their partners' intentions. Thus, Experiment 2b allows us to demonstrate whether individuals are outcome-biased in repeated interactions, where intentions matter more than outcomes.

If joint evaluators in Experiment 2a made decisions to redistribute payoffs, then we expect to replicate the results in the die game, but then find no differences across conditions in individuals' preferences to interact with their partners in the multi-round dictator game. These findings would suggest that joint evaluators were not more outcome-biased than separate evaluators. However, if individuals were reacting to their partners' intentions, then joint evaluators would also be more outcome-biased in their preferences to interact with their partner relative to separate evaluators.

Additionally, whereas participants in Experiment 2a made decisions in hypothetical contexts, participants in Experiment 2b were paired with real partners and made decisions that could have real consequences for their partners.

Method

Participants. We recruited two groups of participants: 1) Player 1s who played the role of die selectors in the die game and the dictator in the multi-round dictator game, and 2) Player

2s—evaluators who made allocations to their partners (Player 1s) based on their partners' die selection and determined whether they wanted to play in the subsequent dictator game.

One hundred and sixty-two individuals (63.6% male, 36.4% female, M_{age} = 35.20 *SD* = 11.06), recruited through Amazon's Mechanical Turk, played the role of the Player 1 (die selector and dictator) in exchange for $0.50. We included several attention checks. For instance, we asked participants to select the last option instead of the correct answer to an arithmetic problem to make sure that they read the instructions carefully. Seven participants did not pass the filter questions and were eliminated from the study automatically.

Two hundred and five participants with a background in economics and finance were recruited from two universities in the northeastern United States (43.7% male, 56.3% female, M_{age} = 21.50 *SD* = 3.25) and played the role of the Player 2 (evaluator) in exchange for a $10 Amazon gift card. All participants passed attention checks and were informed that they would be paid additional money based upon their decisions.

Design and procedure for Player 1s (die selectors and dictators). Participants from Mechanical Turk were all assigned to the role of Player 1 (die selector and dictator) and were informed that they would be playing a die game and an allocation game with a Player 2 (evaluator). These participants were paid $0.50 for their participation with the potential for additional compensation based on the die selected, Player 2's allocations, and their multi-round dictator-game allocations. More specifically, Player 1s (die selectors) were informed they would be paired with Player 2s (evaluators), who would provide either rewards or punishments based on their die selections and outcomes reached. Player 1s read information about the payoff for Die A (the selfish die) and Die B (the fair die), made their die selection, and rolled a virtual die. They were informed that some participants would be randomly selected, and if they were selected,

they would actually receive a $10 endowment, an amount based on the outcome of their die roll, and any additional reward or punishment of up to $9 based on their partner's decision. Thus, Player 1s could earn a minimum of $1 and a maximum of $29 in the die game. Twelve Player 1s with selfish intentions leading to favorable outcomes and another twelve Player 1s with fair intentions leading to undesirable outcomes were selected to be randomly paired with Player 2s and received additional payment based on their partner's decisions.

For the multi-round dictator game, Player 1s (former die selectors now playing the role of dictators) were informed that they would be playing an allocation game that includes five rounds. For each round, they were given an extra $1 and determined how much to send to Player 2 (evaluator). Depending on their decisions, they could keep a minimum of $0 and a maximum of $5 in total. Responses from the same 24 Player 1s selected in the die game were used to determine allocations to Player 2s in the dictator game.

Player 1's selections in the die game (whether they selected the fair die or the selfish die) was significantly correlated with both the average allocation ($r = .19$, $p = .02$) and total allocation ($r = .18$, $p = .03$) in the dictator game, suggesting that fair Player 1s in the die game also allocated more money in the multi-round dictator game (M = $1.59; $SD = 1.11$) relative to selfish partners in the die game (M = $1.16; $SD = 1.19$).

Design and procedure for Player 2s (evaluators). Participants from university subject pools were all assigned to the role of Player 2 and were informed that they would be randomly partnered with Player 1s (die selectors) from another session. All Player 2s (evaluators) were informed that their response could be randomly selected to impact the outcome for their partners.

Die game. Similar to Experiment 2a, all Player 2s (evaluators) were randomly assigned to one of three conditions based on the responses from Player 1s (die selectors): 66 participants

evaluated a fair partner with an unfavorable outcome for the participant, 70 participants evaluated a selfish partner with a favorable outcome for the participant, and 69 participants evaluated both fair and selfish partners with unfavorable and favorable outcomes jointly.

Allocation in die game. We used the same dependent variables as in Experiment 2a. After learning the outcome of the die selections, participants had the opportunity to reward or punish each partner with a maximum of $9. They were told that if they were randomly selected, their decision would impact their partner's (Player 1's) allocation but not their own.

Multi-round dictator game. Player 2s (evaluators) were also informed that they would play another game with the same Player 1—the die selector in the die game. In this game, Player 1 would be given an extra $1 in each round and would determine how much to keep or give to Player 2. Player 2 (the evaluator) was told that Player 1 would receive a total of $5 across five rounds to allocate. Player 2s were informed that if they were randomly selected, they would receive payoffs in this multi-round dictator game based on the outcome of the game.

In the joint evaluation condition, Player 2s (evaluators) were informed that they would play this game with one of their partners, selected at random (the selfish partner or the fair partner in the die game). Player 2s were also informed that both partners determined how much money to allocate to them in each of the five rounds and made these decisions independently. If randomly selected, Player 2s would participate in this multi-round dictator game for real money with their randomly chosen partner.

Value of future interaction. After Player 2s (evaluators) learned the rules of the multi-round dictator game, we elicited their perceived value of playing this game with their partner using the Becker-Degroot-Marschak" (1964) elicitation method. Player 2s were informed that instead of playing this game with their partner, they had the option to receive a certain amount of

money. Player 2s indicated the lowest amount of money between $0 and $5 they were willing to accept for certain in order to take this dollar amount over continuing to play with the same partner. They were informed that a computer would randomly give them an offer of an amount of money between $0 and $5. If the computer offered them an amount greater than the threshold, then Player 2s would keep the amount offered and would not have the opportunity to play in this multi-round dictator game with their partners. However, if the computer offered them an amount less than the threshold, then Player 2s played the dictator game with their partners. Thus, the higher the threshold, the more participants valued playing with their partner. Joint evaluators indicated two values—one for each partner.

Player 2s (evaluators) randomly selected in this game received either the amount that the computer offered or what their partner (Player 1) decided, depending on their decisions and the computer's offered amount. For joint evaluators, one of the partners was randomly selected to play the dictator game, and Player 2s received either the amount that the computer offered or the allocation from the randomly selected partner (dictator).

Taken together, our main dependent variables include (1) whether Player 2s (evaluators) decided to reward or punish Player 1s (die selectors) in the die game ("trembling hand game" in Cushman et al., 2009), (2) the allocation amount in the die game, and (3) the perceived value of participating in a multi-round dictator game with their partner (Player 1).

Results

Allocations in die game. We found a significant interaction between the framing of evaluation and intentions, confirming that participants in the separate-evaluation condition were less outcome-biased, $F(1, 268) = 6.29$, $p = .01$, $\eta_p^2 = .02$. Under joint evaluation, the amount of reward based on evaluation of the fair partner (M = $1.22, SD = 6.62) did not differ significantly

from the selfish partner (*M* = $2.61, *SD* = 5.61), *B* = -1.63, *t*(268) = 1.57, *p* = .12. However, separate evaluators rewarded the partner with fair intentions (*M* = $4.22, *SD* = 5.89) more than they did the partner with selfish intentions, (*M* = $2.17, *SD* = 6.04), *B* = 2.06, *t*(268) = 1.97, *p* = .049, suggesting that separate evaluators were less outcome-biased and paid more attention to intentions than joint evaluators did.

Insert Table 4 and Figure 3 about here

Percent rewarding. We found a marginal interaction between the framing of evaluations and intentions on reward decisions, *B* = .93, Wald χ^2 = 1.89, *p* = .06, odds ratio = 2.55. Joint evaluators rewarded partners with selfish intentions (61.8%) and fair intentions (48.5%) in similar proportions, *B* = .53, Wald χ^2 = 1.53, *p* = .13, odds ratio = 1.70. Separate evaluators also rewarded selfish (57.14%) and fair partners (66.7%) in similar proportions, *B* = -.41, Wald χ^2 = -1.14, *p* = .25, odds ratio = .67.

Percent punishing. We found an interaction between the framing of evaluations and intentions on punishment decisions, *B* = -1.20, Wald χ^2 = -2.05, *p* = .04, odds ratio = .30. Joint evaluators were more likely to punish partners with fair intentions (38.2%) than those with selfish intentions (23.5%), *B* = -.69, Wald χ^2 = -1.84, *p* = .07, odds ratio = .50, whereas separate evaluators punished fair (15.2%) and selfish partners (22.9%) in similar proportions, *B* = .51, Wald χ^2 = 1.14, *p* = .26, odds ratio = 1.66 (see Figure 3).

Value of future interactions with partner. We found that the interaction between the framing of evaluation and intentions is significant, showing that the perceived value of having future interactions with the partner differed based on the mode of evaluation, *F*(1, 264) = 4.79, *p* = .03, η_p^2 = .02.

When participants made decisions jointly, participants valued interacting with partners who had fair intentions in the prior game (M = \$2.57, SD = 1.22) more than selfish partners (M = \$2.24, SD = 1.12), B = .34, $t(264) = 2.35$, $p = .02$. Although separate evaluators also valued playing with fair partners (M = \$3.06, SD = 1.17) more than selfish partners (M = \$2.17, SD = 1.06), B = .89, $t(264) = 4.44$, $p < .001$, that difference was greater in the separate-evaluation condition than the joint-evaluation condition.

Discussion

The results of Experiment 2b provide further support for Hypotheses 1 and 2, showing that separate evaluators factored their partners' intentions into their evaluations and decisions more than joint evaluators did. Not only did separate evaluators punish fair partners less than joint evaluators in a one-shot game, but they also valued having future interactions with their partner more in a multi-round dictator game—a situation where partners' intentions and outcomes are more aligned. These findings rule out the possibility the joint evaluators in the die game were privy to their partners' intentions but chose to ignore their evaluations to satisfy their desire to redistribute earnings evenly. That is, these findings demonstrate that joint evaluators were less likely to attend to information about their partners' intentions, even in contexts where doing so would make them better off in the future.

Experiment 3: Outcome Bias in Decision Making for Others

In Experiment 2b, joint evaluators were more outcome-biased than separate evaluators in a context where individuals were making personal assessments about the value of repeated interactions with either selfish or fair partners. In Experiment 3, we test whether joint evaluators would also be more outcome-biased when making recommendations to others about whether they should have future interactions with selfish or well-intentioned individuals. Additionally,

whereas previous studies employed scenarios in a medical context and economic games, this study examines the impact of outcome information on individuals' evaluations of investment managers in a financial context. We predicted that, third-party recommenders would also be susceptible to the outcome bias when making reward and hiring decisions and that this bias would be stronger in joint-evaluation contexts.

Method

Participants. Three hundred and forty-two individuals (59.9% male, 40.1% female; M_{age} = 35.96, SD = 11.08), recruited through Amazon's Mechanical Turk, participated in a 20-minute online study in exchange for $0.75. Participants were full-time employees who have been working in their current industry for an average of 8.8 years (SD = 8.58). We included attention checks used in Experiment 2b. One participant did not pass through the filter questions and was eliminated from the study automatically.

Design and procedure. Participants were randomly assigned to one of the three conditions: One hundred and sixteen participants evaluated an investment manager with fair intentions that led to an unfavorable outcome, 114 participants evaluated an investment manager with selfish intentions that led to a favorable outcome, and 112 participants evaluated both individuals jointly.

In all conditions, participants were asked to imagine they were giving investment advice to a friend seeking to invest $500,000. In the joint-evaluation condition, participants advised two friends who had different investment managers. These investment managers were responsible for helping clients make sound investment decisions. All participants were given a short introduction with explanations of relevant financial terms and passed a short quiz before participating in the experiment (see Appendix B in the Online Supplement for details).

Participants were informed that the investment manager they were evaluating selected one of two possible funds for a past client: one fund that had higher risk-adjusted returns but offered less compensation to the adviser (Star Trade), or another fund that had lower risk-adjusted returns but offered more compensation to the adviser (Gold Stone). All participants were provided the following financial information about each fund: risk-adjusted returns, volatility, and compensation that these advisers received for investing in the fund.

Star Trade Fund:	Risk-adjusted Return: 0.46
	Volatility: 0.82%
	Compensation to the adviser: 0.1% of initial investment
Gold Stone Fund:	Risk-adjusted Return: 0.35
	Volatility: 3.13%
	Compensation to the adviser: 0.9% of initial investment

The choice of fund signaled the investment managers' intentions to participants. Investment managers who selected the objectively better Star Trade fund for their client sacrificed higher compensation for themselves in service of the client's best interest, whereas those who selected Gold Stone were acting out of self-interest to maximize their own payoff at the expense of the client.

Participants in the separate-evaluation condition were informed that the investment manager they were evaluating chose either the Star Trade or Gold Stone fund. Those evaluating a selfish investment manager learned that the manager chose the Gold Stone Fund and that at the end of the year, Gold Stone outperformed the market and the client made $15,000. In contrast, those evaluating a fair investment manager learned that the manager chose the Star Trade Fund, which underperformed the market and led the client to lose $15,000. (See the Online Supplement for the full protocol.) Therefore, half of the participants in the separate-evaluation condition evaluated an investment manager who gave clients favorable outcomes despite having selfish

intentions, while the other half evaluated a manager who offered unfavorable outcomes to clients despite fair intentions.

In the joint-evaluation condition, participants evaluated both investment managers (the fair manager whose decision resulted in an unfavorable outcome for the client and the selfish manager whose decision resulted in a favorable outcome for the client) and were informed that the two managers independently determined which fund to select. Both managers' selections and the outcome of their decisions were presented simultaneously to the managers.

Dependent measures. Next, participants were informed that their friend was seeking their recommendation regarding how much of $500,000 should be actively managed by their respective investment manager and how much of it should be passively invested in the S&P 500. Participants were given a slider ($0 - $500,000) to indicate how much of the total $500,000 amount they thought their friend should allocate to the investment manager. As they indicated their decision on the slider, a second slider below automatically adjusted the remaining amount that would go to S&P 500, such that 100% of the money would be invested.

In addition, participants rated the investment manager's intentions on the following two items: the extent to which the manager acted in the client's best interest (1 = *not at all*, 7 = *very much*) and how they would describe the investment manager's intentions (1 = *extremely bad*, 7 = *extremely good*) (α = .92). Additionally, participants rated the extent to which they were paying attention to the manager's fund selection relative to return for the past client. The slider had a range from -100 ("only investment manager's past investment choice") to 100 ("only investment manager's past return for client"), but the values were not visible to the participants. Participants also indicated the extent to which they liked and felt close to the manager (α = .93; Rudman,

1998) and indicated their own likelihood of hiring the investment manager (1 = *not at all*, 7 = *very much*). Finally, participants answered demographic questions.

Results

For brevity, results on self-reported attentiveness to intentions, liking and hiring are reported in the Online Supplement.

Allocation recommendation. We found a significant interaction between framing of evaluation and intentions, $F(1, 450) = 13.09$, $p < .001$, $\eta_p^2 = .03$. Joint evaluators recommended allocating more to the selfish investment manager (M = \$249,422.12, SD = 112,154.63) than to the fair manager (M = \$181,856.48 SD = 112,178.35), B = 67,565.63, $t(450) = 6.04$, $p < .001$. However, separate evaluators recommended allocating similar amounts of money to the fair investment manager (M = \$177,287.94, SD = 115,419.58) and the selfish one (M = \$175,333.42, SD = 131,626), B = -1,954.52, $t(450) = .13$, $p = .90$, suggesting that separate evaluators were less outcome-biased than joint evaluators.

Intentions. We found a significant interaction between outcome and mode of evaluation, $F(1, 450) = 12.36$, $p < .001$, $\eta_p^2 = .02$, suggesting that managers with favorable and unfavorable outcomes were perceived differently under separate and joint evaluations. Joint evaluators rated fair-intentioned advisers (M = 5.01, SD = 1.31) and selfish advisers (M = 4.69, SD = 1.46) similarly, B = .32, $t(450) = 1.66$, $p = .09$. However, separate evaluators rated fair advisers' intentions higher (M = 5.06, SD = 1.28) than the intentions of selfish advisers (M =3.77, SD=1.72), B = -1.28, $t(450) = 6.69$, $p < .001$.

Insert Table 5 about here

Attention to manager's fund selection relative to the past client's returns. We found a significant interaction between the outcome and evaluation mode, $F(1, 450) = 5.64$, $p = .02$, $\eta_p^2 = .02$. Joint evaluators indicated that they were paying attention to outcomes when they evaluated both selfish investment managers ($M = 16.54$, $SD = 52.99$) and fair managers ($M = 13.78$, $SD = 51.51$), $B = 2.77$, $t(450) = .76$, $p = .45$. In contrast, when participants made their evaluations separately, attention ratings for the manager with selfish intentions ($M = -3.85$, $SD = 54.54$) were significantly different from those for the manager with fair intentions ($M = 11.78$, $SD = 48.31$), $B = -15.64$, $t(450) = 2.29$, $p = .02$. We also found a marginally significant main effect of evaluation mode, as separate evaluators paid less attention to outcomes (and more attention to intentions) ($M = 4.04$, $SD = 51.97$) than joint evaluators did ($M = 15.16$, $SD = 52.16$), $F(1, 450) = 3.86$, $p = .05$, $\eta_p^2 = .009$.

We tested for mediated moderation to examine whether perceptions of the partner's intentions explain effect of the interaction between evaluation mode and outcome on allocation decisions (Baron & Kenny, 1986). When controlling for perceptions of the partner's intentions, the interaction between evaluation mode and outcome was significantly reduced (from $B = 105.16$, $t = 3.15$, $p = .002$ to $B = 55.97$, $t = 1.82$, $p = .07$), and perceived intentions predicted allocation decisions ($B = 51.34$, $t = 9.76$, $p < .001$). A bootstrap analysis indicated that the 95% bias-corrected confidence interval for the size of the indirect effect excluded zero for separate evaluators (-90.83, -44.94), but included zero for joint evaluators (-36.13, 2.64), revealing that perceptions of the partner's intentions mediated the relationship between outcome and allocation decision for separate evaluators but not for joint evaluators (MacKinnon et al., 2007).

Discussion

The results of Experiment 3 demonstrate that joint evaluation exacerbates the outcome bias in organizational contexts, even when evaluators do not receive any direct benefits from rewarding outcomes. Furthermore, self-reports of evaluators demonstrate that joint evaluators were less attentive to their partners' intentions, explaining our effect that joint evaluation exacerbates the outcome bias.

Experiment 4: Lack of Consideration of Intentions Under Joint Evaluation

Building on the results of Experiments 1-3, Experiment 4 sought to reduce the outcome bias through interventions requiring individuals not only to make judgments about their partners before learning their outcomes, but also to make preliminary reward decisions before outcomes are revealed. Beyond testing a potential intervention to mitigate the outcome bias, Experiment 4 also serves as a test of the mechanism underlying our effect. Whereas Experiment 3 tested the role of perceived intentions via mediation, we test its role via moderation (Spencer, Zanna, & Fong, 2005) by directly manipulating the extent to which intentions are salient.

Specifically, half of the joint and separate evaluators rewarded their partner at the end of the study, as in Experiments 2a and 2b—after they discovered their partner's die selection and the outcome of the die roll. The other half of joint and separate evaluators not only made allocation decisions at the end of the study, but also made preliminary allocation decisions for their partners after they discovered their partner's intentions but before they discovered the outcome of the die rolls. Participants made decisions that could have real financial consequences for their partners. Because joint evaluators were less attentive to intentions, we predicted that making preliminary reward decisions would be particularly helpful in reducing reliance on outcomes for joint-evaluators.

Method

Participants. Five hundred forty-two participants (58% male, 42% female; M_{age} = 30.07, *SD* = 10.71) adopting the role of Player 2s were recruited through Amazon's Mechanical Turk to participate in a 20-minute online study in exchange for $0.75. Player 2s were asked to reward or punish their partner, selected from a separate group of 224 Mechanical Turk participants. These partners were paid $0.25 with the potential for additional compensation based on the die selected and the allocation of their assigned Player 2 (52% male, 48% female; M_{age} = 32.33, *SD* = 11.83).

Design and procedure. Evaluators were randomly assigned to one of six conditions in a 2 (preliminary allocation vs. no preliminary allocation) x 3 (separate evaluation of individuals with generous intentions leading to unfavorable outcomes, separate evaluation of those with selfish intentions leading to favorable outcomes, and joint evaluation of both types of individuals) study design.[2]

As in Experiment 2, separate evaluators in the role of Player 2 were randomly paired with one of 34 previous MTurk participants (from the pool of 224 participants recruited to be Player 2's partners). Those making preliminary decisions were asked to provide a reward and punishment decision and then rate the intentionality and generosity of each partner before finding out the outcome of the die selection. Upon learning the final outcomes, participants were reminded of their initial reward and punishment decisions for each partner and then asked to make final reward and punishment decisions.

In contrast, those who did not make a preliminary allocation decision followed a similar protocol to that described in Experiment 2 and then answered questions about the intentionality and generosity of their partner after learning the final outcome. At the end of the study, all participants rated their partner and answered demographic questions.

Participants were informed that if they were randomly selected, their decisions would have real consequences for their partner. More specifically, five percent of evaluators (27 Player 2s) were randomly selected to receive payment based on the outcome of the die roll (either $0 or $10) and to issue the rewards or punishments they allocated to their partner. Those asked to make preliminary allocations were informed that if they were randomly selected, there was a 25% chance that their preliminary reward or punishment decision would be meted out to their partner and a 75% chance that their final reward or punishment decision would be enacted.[3] Joint evaluators were instructed that if they were randomly selected, both partners would receive their respective rewards and punishments. Evaluators (Player 2s) did not incur any financial costs or benefits based on their reward or punishment decisions. That is, they did not need to give away their own money to reward their partners, and they could not keep any money gained from punishing their partners.

Dependent measures. Participants were asked to reward or punish die selectors up to a maximum of $9 after finding out the outcome of the die selection. Those asked to make preliminary allocations could reward or punish die selectors up to a maximum of $9 immediately after finding out the partner's die selection. Participants making preliminary allocations also rated the generosity (1 = extremely selfish, 7 = extremely generous) of their partner before finding out the outcome of the die roll. Those who were not asked to make preliminary allocations rated generosity after making final reward and punishment decisions. At the end of the study, all participants rated items that reflected the extent to which they felt close to their partner. These seven items were adapted from a social-attraction scale ($\alpha = .90$) (Rudman, 1998). They also self-reported the degree to which they remembered their partner's die selection and outcomes well (1 = *not at all*, 7 = *extremely*).

Results

We present initial and final allocation decisions below; details about perceptions of the partner's generosity and liking are reported in Appendix C of the Online Supplement.

Allocations. Eliciting preliminary allocations reduced the outcome bias under joint-evaluation contexts but not in separate-evaluation contexts. We found a significant three-way interaction across the framing of evaluation (joint vs. separate), type of outcome (unintentionally favorable or unfavorable), and preliminary reward, $F(1, 705) = 6.81$, $p = .009$, $\eta_p^2 = .009$ (see Figure 4).

There does not appear to be evidence that making preliminary allocations reduces reliance on outcomes in separate-evaluation contexts, as there is not a significant interaction between the type of outcome evaluated and whether participants made preliminary allocations, $B = .52$, $t(705) = .48$, $p = .63$. Under separate evaluation, the amount of rewards did not differ significantly between selfish partners and fair partners for both cases when participants were asked to make preliminary decisions ($M_{selfish} = \$3.32$, $SD_{selfish} = 4.75$, $M_{fair} = \$3.78$, $SD_{fair} = 5.20$), $B = .46$, $t(705) = .60$, $p = .58$, and when participants were not given this reminder ($M_{selfish} = \$2.79$, $SD_{selfish} = 5.05$, $M_{fair} = \$3.76$, $SD_{fair} = 6.17$), $B = .98$, $t(705) = 1.31$, $p = .19$.

Insert Table 6 and Figure 4 about here

However, making preliminary allocations mitigated participants' reliance on outcomes in joint-evaluation contexts, as we found a significant interaction between the type of outcome evaluated and whether participants made preliminary allocations, $B = -3.42$, $t(705) = -3.24$, $p = .001$. Joint-evaluation rewards were higher for selfish partners ($M = \$2.82$, $SD = 3.76$) than for fair partners ($M = -\$0.46$, $SD = 5.22$) when participants were not asked to make a

preliminary allocation, B = -3.29, $t(705) = 4.16$, $p < .001$. However, these rewards were not significantly different from one another when participants were asked to make a preliminary decision ($M_{selfish}$ = \$2.51, $SD_{selfish}$ = 4.94, M_{fair} = \$2.65, SD_{fair} = 5.45), B = .14, $t(705)$ = .20, p = .85.

Percent rewarding. We found a significant three-way interaction across the framing of evaluation (joint vs. separate), type of outcome (unintentionally favorable or unfavorable), and preliminary decision, B = -1.35, Wald χ^2 = -2.03, p = .04, odds ratio = .26. When not asked to make a preliminary decision, joint evaluators were more likely to reward selfish partners (74%) than fair partners (39%), B = 1.50, Wald χ^2 = 4.28, $p < .001$, odds ratio = 4.47, whereas separate evaluators rewarded selfish partners (71%) and fair partners (69%) in similar proportions, B = .12, Wald χ^2 = .39, p = .69, odds ratio = 1.13.

However, after being asked to make a preliminary decision, joint and separate evaluators acted similarly. Joint evaluators rewarded selfish partners (73%) and fair partners (68%) in similar proportions, B = .24, Wald χ^2 = .78, p = .44, odds ratio = 1.27, just as separate evaluators rewarded selfish partners (77%) and fair partners (73%) in similar proportions, B = .22, Wald χ^2 = .62, p = .54, odds ratio = 1.24.

Percent punishing. We also found a significant three-way interaction across the framing of evaluation (joint vs. separate), type of outcome (unintentionally favorable or unfavorable), and preliminary decision, B = 1.68, Wald χ^2 = 2.10, p = .04, odds ratio = 5.34. When not asked to make a preliminary decision, joint evaluators were more likely to punish fair partners (48%) than selfish partners (8%), B = -2.39, Wald χ^2 = -4.96, $p < .001$, odds ratio = .09, whereas separate evaluators punished fair partners (25%) and selfish partners (18%) similarly, B = -.43, Wald χ^2 = -1.20, p = .23, odds ratio = .65.

However, after being asked to make a preliminary decision, joint evaluators behaved very similarly to separate evaluators. Joint evaluators punished fair partners (23%) in similar proportion to selfish partners (19%), $B = -.24$, Wald $\chi^2 = -.69$, $p = .49$, odds ratio = .79, just as separate evaluators also punished fair partners (18%) in similar proportion to selfish partners (18%), $B = .05$, Wald $\chi^2 = .13$, $p = .90$, odds ratio = 1.05.

Initial allocations. An analysis of initial allocations revealed a main effect of the partner's intentions, such that participants rewarded fair partners ($M = 5.45$, $SD = 3.94$) more than they did selfish partners ($M = -1.33$, $SD = 6.20$), $F(3, 367) = 59.41$, $p < .001$, $\eta_p^2 = .33$. As expected, this analysis did not reveal a significant main effect of the type of evaluation (joint vs. separate) or an interaction between the type of evaluation and intentions, Fs < .72, ps > .19, suggesting that the type of evaluation did not appear to affect initial allocations made.

Discussion

When participants were not asked to make preliminary allocations, we obtained the same findings as in Experiments 1–3 . However, making preliminary allocations before discovering the outcomes effectively mitigated the outcome bias arising from complex joint-evaluation contexts. As expected, we replicated findings in Experiments 1 and 2 when participants were not asked to make preliminary reward decisions; however, when they made preliminary reward decisions, our findings suggest that interventions requiring individuals to make judgments prior to realizing outcomes reduces the outcome bias in joint-evaluation contexts.

Furthermore, our evidence suggests that although participants had similar initial judgments of their partners' intentionality in both joint- and separate-evaluation contexts, presentation of outcomes overrides individuals' initial impressions in joint-evaluation contexts but not separate-evaluation contexts. We note that our intervention did not affect final allocations

for separate evaluators, suggesting that separate evaluation does not impair individuals' ability to factor initial intentions into their final judgments.

Beyond providing more information about the mechanism underlying the finding that joint comparisons can increase the outcome bias, these findings also have practical relevance: requiring individuals to make judgments prior to realizing outcomes appears to be an effective intervention for reducing the outcome bias in joint-evaluation contexts.

General Discussion and Conclusions

Individuals not only neglect information about procedures and decision quality when facing information about outcomes (Adams, 1965; Deutsch, 1985; Mazzocco, Alicke, & Davis 2004; Robbennolt, 2000), but they also neglect information about intentions, particularly when making joint comparisons. The present studies specifically investigate how different modes of evaluation affect bias against intentions and provide a strategy for mitigating this bias. Our findings suggest that joint-evaluation contexts that present multiple alternatives exacerbate the outcome bias, making important attributes, such as intentions, more difficult to consider. This paper has demonstrated that making intentions more salient is an effective method of mitigating the outcome bias, particularly in joint-evaluation contexts.

Our findings extend prior research on the outcome bias by examining the role of intentions and by examining the impact of different evaluation modes on decision evaluation in a variety of contexts, including medical and financial decision making as well as economic games. Experiments 1-3 found that information about intentions was more influential when participants evaluated one option at a time as compared to when they evaluated multiple decisions simultaneously. In fact, we found that joint evaluation exacerbates the outcome bias, whereas separate evaluation reduces it. Given that many hiring and promotion decisions are made jointly,

Experiment 4 shows that managers could overcome the outcome bias by making their initial evaluations of employees' decisions before the outcomes of those decisions are fully known.

These findings that joint evaluation exacerbates the outcome bias seemingly contradict advice based on prior literature that prescribes joint evaluation as a tool for mitigating decision-making biases (Bazerman et al., 2011; Bohnet et al., in press; Gino et al., 2010). Joint evaluation in prior research provided a comparison point that highlighted the key differences on *one* relevant attribute that was easily ignored in separate-evaluation contexts (Bazerman, Loewenstein & White, 1992; Nowlis & Simonson, 1997; Kahneman et al., 1993; Hsee et al., 1999; Gino et al., 2010). For example, in Study 5 of Gino, Shu, and Bazerman (2010), participants rated the ethicality of a physician's decisions in which the only dimension that was varied was the outcome of the physician's actions. In these situations, joint evaluators were less biased than separate evaluators.

However, joint evaluation can be harmful when (1) attributes vary on *both intentions and outcomes* and (2) outcomes are more salient than intentions. In our case, intentions were more difficult to infer (e.g., based on selection of Die A or Die B), whereas outcomes were much more salient (e.g., allocations from a die roll). The results of Experiment 4 provide empirical support for our hypothesis of the mechanism underlying these results, as making intentions more salient led joint evaluators to be no more outcome-biased than separate evaluators. These finding suggest that these complex joint evaluations led participants to ignore information about intentions unintentionally.

Taken together, our findings bridge decision-making research on the outcome bias (Baron & Hershey, 1988; Hastie & Dawes, 2001; Gino et al., 2012) and management research on organizational justice (Leventhal, 1980; Thibaut & Walker, 1975). Prior decision-making and

psychology research has demonstrated that individuals tend to take outcome information into account even if it does not provide any additional information about decision quality (Allison, Mackie & Messick, 1996; Baron & Hershey, 1988). Independently, management scholars have also focused on the role of outcomes and procedures in predicting individuals' sense of fairness and justice in a given environment (Brockner & Wiesenfeld, 1996; Cropanzano & Greenberg, 1997; Deutsch, 1985; Leventhal, 1980; Lind & Tyler, 1988; Thibaut & Walker, 1975).

Our findings integrate these literatures by investigating a relevant factor that influences decision quality and the fairness of procedures that determine outcomes: *intentions.* Intentions, processes, and outcomes vary independently from one another, and future research is needed to understand how all three factors interact to influence individuals' evaluations. Furthermore, the study of intentions and processes has thus far focused on the ethical and harmful impacts of neglecting information about "good" or "bad" intentions and processes. However, information about intentions and processes is rarely straightforward (either good or bad), and actions based on this information have wider implications beyond harmful or unfair evaluations. Future research is needed to study how individuals react to more complex information about intentions and processes and how this knowledge may also influence evaluations and interpersonal dynamics in organizations.

Because research has demonstrated that joint evaluation often encourages deliberate decision making and reduces decision-making biases (Bazerman et al., 2011; Gino et al., 2010), more research is needed to identify the conditions under which joint evaluations can exacerbate other biases and the boundary conditions that determine when joint evaluations prove unhelpful or harmful. For example, does joint evaluation exacerbate biases in other contexts—beyond the comparison of intentions and outcomes—where attributes vary on multiple categories?

Additionally, future research is needed to fine-tune our understanding of the mechanisms driving the the outcome bias under joint evaluation. We propose that joint evaluation can make some attributes—such as intentions—less salient and thus more difficult to consider, but different mechanisms may explain these findings. This knowledge can help researchers and practitioners more fully understand the benefits and risks of each mode of evaluation.

Conclusion

People regularly make and receive evaluations. Although reliance on outcomes in these evaluations leads to suboptimal outcomes (Bazerman & Moore, 2013; Hastie & Dawes, 2001), little research has explored strategies aimed at reducing the failure to examine intentions. Investigating the effects of different modes of evaluation on the outcome bias when the decision maker's intentions are known, the present research shows that joint evaluation can exacerbate this bias. These findings offer a new perspective on the distinction between joint and separate evaluation and the conditions under which each improves decision making.

References

Adams, J. S. (1965). Inequity in social exchange. *Advances in experimental social psychology*, *2*, 267-299.

Alicke, M. D., & Davis, T. L. (1989). The role of a posteriori victim information in judgments of blame and sanction. *Journal of Experimental Social Psychology*,*25*(4), 362-377.

Allison, S. T., Mackie, D. M., & Messick, D. M. (1996). Outcome biases in social perception: Implications for dispositional inference, attitude change, stereotyping, and social behavior. In M. P. Zanna (Eds.), *Advances in experimental social psychology, 28,* 53-93. New York: Academic.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54,* 569-579.

Bazerman, M. H., Gino, F., Shu, L. L., & Tsay, C. (2011). Joint evaluation as a real world tool for managing emotional assessment of morality. *Emotion Review, 3*(3), 290-292.

Bazerman, M. H., Loewenstein, G. F., & White, S. B. (1992). Reversals of preference in allocation decisions: Judging an alternative vs. choosing among alternatives. *Administrative Science Quarterly, 18,* 220–240.

Bazerman, M. H., & Moore, D.A. (2013). *Judgment in Managerial Decision Making*, 8th ed. New York: Wiley.

Bazerman, M. H., Moore, D.A., Tenbrunsel, A. E., Wade-Benzoni, K. A., & Blount, S. (1999). Explaining how preferences change across joint versus separate evaluation. *Journal of Economic Behavior & Organization, 39*(1) 41-58.

Bazerman, M. H., Schroth, H. A., Shah, P. P., Diekmann, K. A., & Tenbrunsel, A. E. (1994). The inconsistent role of social comparison and procedural justice in reactions to

hypothetical job descriptions: Implications for job acceptance decisions. *Organizational Behavior and Human Decision Processes, 60,* 326–352.

Bazerman, M.H., Tenbrunsel, A. E., & Wade-Benzoni, K. A. (1998). Negotiating with yourself and losing: Understanding and managing competing internal preferences. *Academy of Management Review, 23,* 225–241.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, *9*(3), 226-232.

Bohnet, I., van Geen, A., & Bazerman, M.H. (in press). When performance trumps gender bias: Joint versus separate evaluation. *Management Science*.

Brockner, J., & Wiesenfeld, B.M. (1996). An integrative framework for explaining reactions to decisions: Interactive effects of outcomes and procedures. *Psychological Bulletin, 120,*189-208.

Charness, G., & Levine, D. I. (2007). Intention and stochastic outcomes: An experimental study. *Economic Journal*, *117*, 1051–1072.

Cropanzano, R., & Greenberg, J. (1997). Progress in organizational justice: Tunneling through the maze. *International review of Industrial and Organizational Psychology, 12,* 317-372.

Cushman F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition. 108*, 353–380.

Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a “trembling hand” game. *PloS one*, *5*, 1-7.

Dennett, D. (1987). True believers. *The intentional stance*. Cambridge: MIT.

Deutsch, M. (1985). *Distributive Justice: A social psychological perspective*. New Haven: Yale University Press.

Fiske, S. T. (1989). Examining the role of intent: Toward understanding its role in stereotyping and prejudice. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought,* 253-283. New York: Guilford Press.

Gerstenberg, T., Lagnado, D. A., & Kareev, Y. (2010). The dice are cast: The role of intended versus actual contributions in responsibility attribution. In S. Ohlsson & R. Catrambone (Eds.), *Cognition in flux*: *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, 1697–1702. Austin, TX: Cognitive Science Society.

Gino, F., Moore, D. A., & Bazerman, M. H. (2012). *No harm, no foul: The outcome bias in ethical judgments*. Working paper.

Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless + Harmless = Blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human Decision Processes, 111*(2), 102-115.

Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology*, *79*, 691– 701.

Handgraaf, M. J., Dijk, E. V., Wilke, H. A., & Vermunt, R. C. (2004). Evaluability of outcomes in ultimatum bargaining. *Organizational Behavior and Human Decision Processes*, *95*(1), 97-106.

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks, CA: Sage.

Hawkins, S., & Hastie, R. (1990). Hindsight biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107,* 311–327.

Heider, F. (1958). *Interpersonal relations*. New York: Wiley.

Hsee, C.K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67,* 247–257.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin, 125*(5), 576-590.

Jones, E. E., & Nisbett, R. E. (1971). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior*, 79-94. Morristown, NJ: General Learning Press.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136-153.

Kahneman, D., Ritov, I., Jacowitz, K. E., & Grant, P. (1993). Stated willingness to pay for public goods: A psychological perspective. *Psychological Science*, *4*(5), 310-315.

Kruger, J., & Gilovich, T. (2004). Actions, intentions, and self-assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin, 30*(3), 328-339.

Leventhal, G. S. (1980). *What should be done with equity theory?* 27-55. Springer US.

Levitt, A., Dwyer, P. (2002). *Take on the street*. New York: Pantheon.

Lind, E. A., & Lissak, R. I. (1985). Apparent impropriety and procedural fairness judgments. *Journal of Experimental Social Psychology*, *13*, 338-350.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Springer.

Lowe, C. A., & Medway, F. J. (1976). Effects of valence, severity, arid relevance on responsibility and dispositional attribution. *Journal of Personality*, *44*(3), 518-538.

Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology, 26*, 131-146.

Messick, D. M., & Sentis, K. (1979). Fairness and preference. *Journal of Experimental Social Psychology*, *15*(4), 418-434.

McFarlin, D. B., & Sweeney, P. D. (1992). Distributive and procedural justice as predictors of satisfaction with personal and organizational outcomes. *Academy of Management Journal*, *35*, 626– 637.

Milkman, K.L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved. *Perspectives on Psychological Science, 4*(4), 379-383.

Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research, 17*(2), 189-202.

Moore, D., Tetlock, P., Tanlu, L., & Bazerman, M. H. (2006). Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review, 31*(1), 1-20.

Nowlis, S. M., & Simonson, I. (1997). Attribute-task compatibility as a determinant of consumer preference reversals. *Journal of marketing research*, 205-218.

Pizarro D., Uhlmann, E., & Bloom, P. (2003). Causal deviance and the attribution of moral responsibility**.** *Journal of Experimental Social Psychology, 39*(6), 653-660.

Robbennolt, J. K. (2000). Outcome severity and judgments of "responsibility": A meta-analytic review. *Journal of Applied Social Psychology, 30,* 2575-2609.

Rudman, L.A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counter-stereotypical impression management. *Journal of Personality and Social Psychology, 74*(3), 629-645.

Schroeder, D. A., & Linder, D. E. (1976). Effects of actor's causal role, outcome severity, and knowledge of prior accidents upon attributions of responsibility. *Journal of Experimental Social Psychology*, *12*(4), 340-356.

Shultz, T. R., & Wells, D. (1985). Judging the intentionality of action-outcomes. *Developmental Psychology*, *21*(1), 83-89.

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness.* New York: Springer-Verlag.

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, 845–851.

Stanca, L. (2010). How to be kind: Outcomes versus intentions as determinants of fairness. *Economic Letters*, *106,* 19–2.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*, 1141-1160, Hillsdale, NJ; New York, NY: L. Erlbaum Associates.

Wang, L., & Murnighan, J. K. (2013). The generalist bias. *Organizational Behavior and Human Decision Processes, 120* (1), 47-61.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct.* New York: Guilford Press.

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2004). “Is 28% good or bad?” Evaluability and preference reversals in health care decisions. *Medical Decision Making*, *24*(2), 142-148.

Tables

Table 1

Descriptive statistics by condition for the dependent measures in Experiment 1.

Standard deviations are reported in parentheses.

		Doctor Ratings
Separate	Selfish Intentions, Favorable Outcome	3.67 (1.00)
	Fair Intentions, Unfavorable Outcome	3.33 (1.18)
Joint	Selfish Intentions, Favorable Outcome	4.14 (0.89)
	Fair Intentions, Unfavorable Outcome	2.79 (1.27)

Table 2

Payoff distribution of Die A and Die B in Experiment 2.

Payoff for Die A (Selfish Die)	**1**	**2**	**3**	**4**	**5**	**6**
Evaluator (Player 2)	$0	$0	$0	$0	$5	$10
Die Selector (Player 1)	$10	$10	$10	$10	$5	$0
Payoff for Die B (Fair Die)						
Evaluator (Player 2)	$5	$5	$5	$5	$0	$10
Die Selector (Player 1)	$5	$5	$5	$5	$10	$0

Table 3

Descriptive statistics by condition for the rewards and punishments measured in Experiment 2a. Standard deviations are reported in parentheses for the continuous variable, and exact proportions are provided in parentheses for binary variables.

		Allocation to Partner	Percent Reward	Percent Punishment
Separate	Selfish Intentions, Favorable Outcome	$2.44 (5.82)	57% (57/100)	24% (24/100)
	Fair Intentions, Unfavorable Outcome	$2.09 (6.43)	51% (50/99)	29% (29/99)
Joint	Selfish Intentions, Favorable Outcome	$4.20 (4.54)	72% (73/101)	10% (10/101)
	Fair Intentions, Unfavorable Outcome	$0.39 (5.64)	41% (41/101)	33% (33/101)

Table 4

Descriptive statistics by condition for the rewards and punishments measured in Experiment 2b. Standard deviations are reported in parentheses for continuous variables, and exact proportions are provided in parentheses for binary variables.

		Allocation to Partner	Percent Reward	Percent Punishment	Value of Playing with Partner
Separate	Selfish Intentions, Favorable Outcome	$2.17 (6.04)	57.14% (40/70)	22.85% (16/70)	$2.17 (1.06)
	Fair Intentions, Unfavorable Outcome	$4.22 (5.89)	66.7% (44/66)	15.15% (10/66)	$3.06 (1.17)
Joint	Selfish Intentions, Favorable Outcome	$2.61 (5.61)	61.8% (42/68)	23.5% (16/68)	$2.24 (1.12)
	Fair Intentions, Unfavorable Outcome	$1.22 (6.62)	48.5% (33/68)	38.2% (26/68)	$2.57 (1.22)

Table 5

Descriptive statistics by condition for the reward and punishment, best interest, liking, and promotion measured in Experiment 3. Standard deviations are reported in parentheses.

		Allocation to Manager	Intentions	Attention
Separate	Selfish Intentions, Favorable Outcome	$175,333.42 (131,626.002)	3.77 (1.72)	3.85 (54.54)
	Fair Intentions, Unfavorable Outcome	$177,287.94 (115419.58)	5.06 (1.28)	11.78 (48.31)
Joint	Selfish Intentions, Favorable Outcome	$251,669.16 (110,101.56)	4.69 (1.46)	16.54 (52.99)
	Fair Intentions, Unfavorable Outcome	$178990.32 (108,489.42)	5.01 (1.31)	13.78 (51.51)

Table 6

Regression results for allocation, reward, and punishment decisions in Study 4.

Variable	Allocation	Reward	Punish
	(1)	(2)	(3)
Evaluation mode *(0=separate, 1=joint)*	-4.23***	-1.23***	1.01***
Intentions *(0=selfish, 1=fair)*	-.98	.12	-.43
Preliminary decision *(0=control, 1=treatment)*	.02	.21	-.46
Evaluation mode x Intentions	4.26***	1.37**	-1.97**
Evaluation mode x preliminary decision	3.10**	1.00*	.17
Intentions x preliminary decision	.52	.09	.48
Evaluation mode x preliminary decision x Intentions	-3.94**	-1.35*	1.68*
Intercept	3.76	.79	-1.09***
Model	Linear	Logistic	Logistic

The table reports unstandardized coefficients for all regressions.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Figures

Figure 1

Physician ratings by condition for Experiment 1.

Error bars reflect ± 1 standard error from the mean.

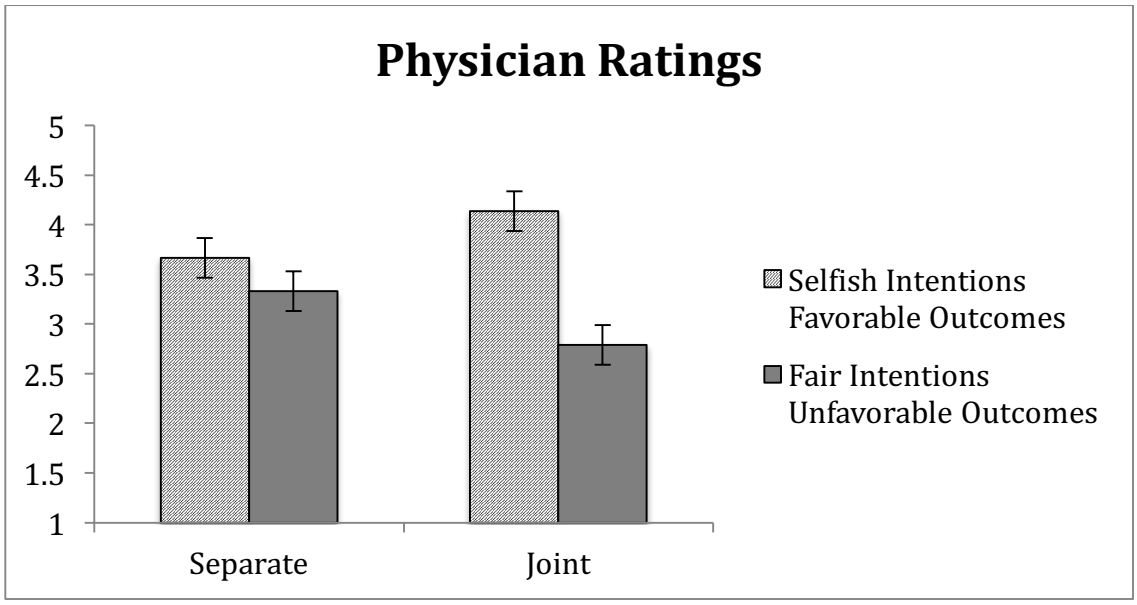


Figure 2

Reward and punishment decisions by condition for Experiment 2a.

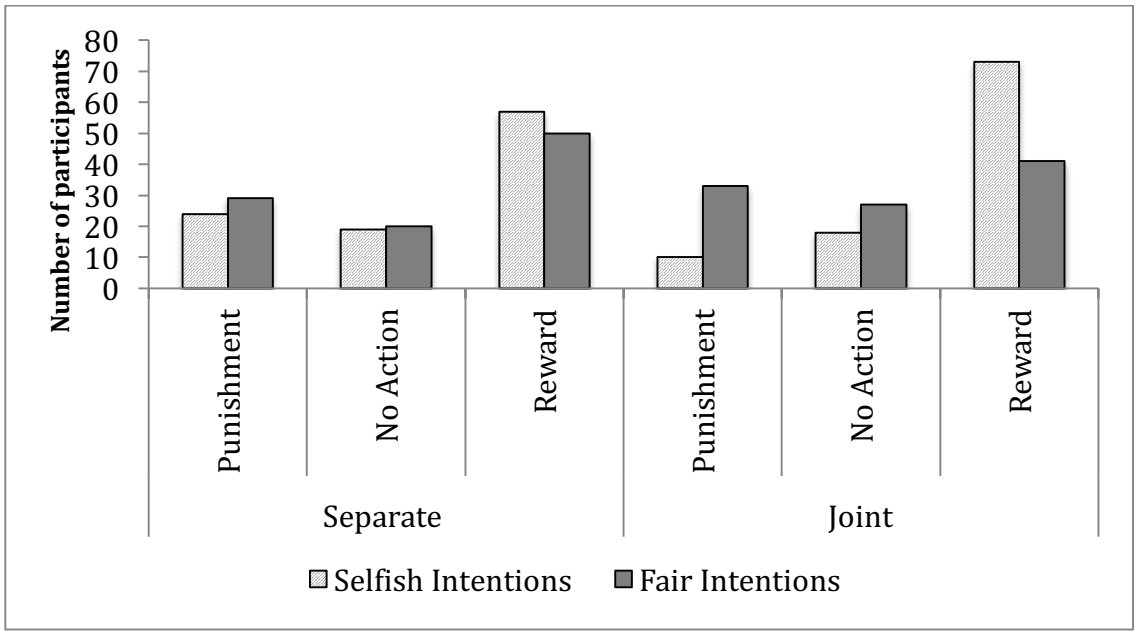


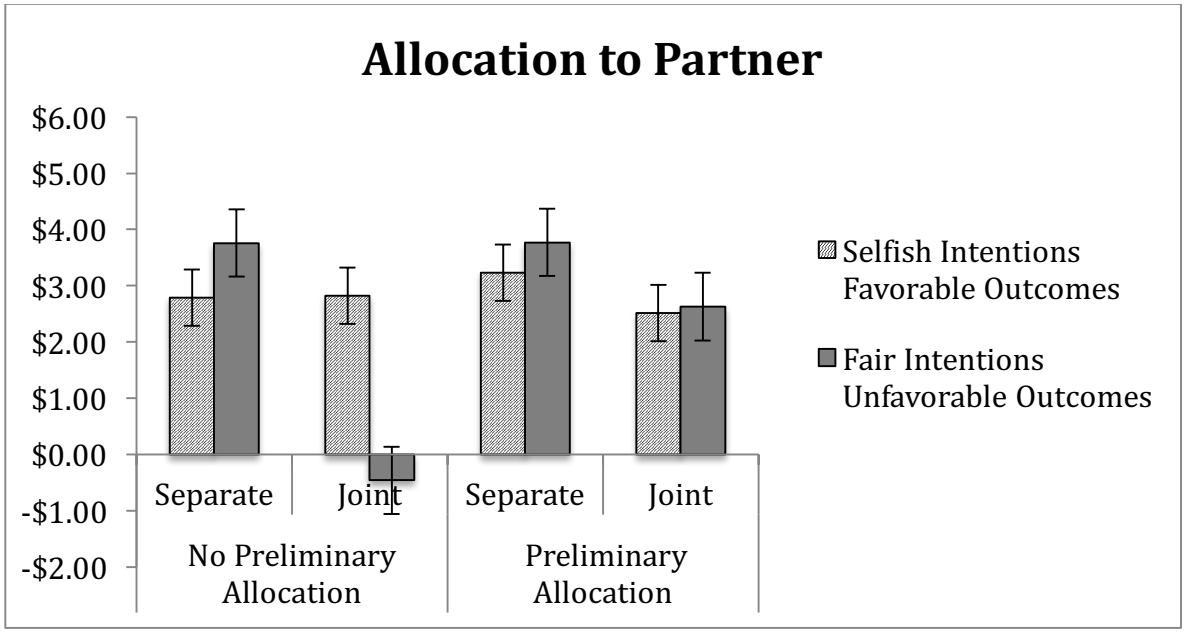
Figure 3

Reward and punishment decisions by condition for Experiment 2b.

Figure 4

Allocation to partner by condition for Experiment 4.

Error bars reflect ± 1 standard error from the mean.



Footnotes

[1] For our analyses, we dummy coded our independent variables.

[2] In Study 4, from a total sample size of N = 542, 99 participants made separate evaluations of fair partners, but were not asked to allocate a preliminary reward; 85 participants made separate evaluations of fair partners and were asked to make a preliminary reward; 94 participants were randomly assigned to separate evaluation of selfish partners, but were not asked to allocate a preliminary reward; 87 participants were randomly assigned to separate evaluation of selfish partners and made a preliminary reward decision; 77 participants made joint evaluations but did not make a preliminary reward, and 100 participants made joint evaluations with preliminary reward.

[3] For those in the preliminary intentions condition, we set the incentive structure (25% preliminary and 75% final) so that participants were incented 1) to pay attention to the information about their partners' intentions and 2) to spend more time making the final decision based on all of the information available (the partner's intentions based on the die selected and outcomes based on the result of the die roll).