

09-020

**Nameless + Harmless =  
Blameless:  
When Seemingly Irrelevant  
Factors Influence Judgment  
of (Un)ethical Behavior**

**Francesca Gino  
Lisa L. Shu  
Max H. Bazerman**

Copyright © 2008 by Francesca Gino, Lisa L. Shu, and Max H. Bazerman

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Running head: NAMELESS + HARMLESS = BLAMELESS

Nameless + Harmless = Blameless:

When Seemingly Irrelevant Factors Influence Judgment of (Un)ethical Behavior

Francesca Gino,<sup>1</sup> Lisa L. Shu,<sup>2</sup> and Max H. Bazerman<sup>2</sup>

<sup>1</sup> *Kenan-Flagler Business School, University of North Carolina*

<sup>2</sup> *Harvard Business School, Harvard University*

The authors thank the Center for Behavioral Decision Research at Carnegie Mellon University for its support. We thank Kimberly Ling, Lamar Pierce, and Zachariah Sharek for their thoughtful comments on earlier versions of this paper. Please address correspondence to fgin@unc.edu.

### Abstract

People often make judgments about the ethicality of others' behaviors and then decide how harshly to punish such behaviors. When they make these judgments and decisions, sometimes the victims of the unethical behavior are identifiable, and sometimes they are not. In addition, in our uncertain world, sometimes an unethical action causes harm, and sometimes it does not. We argue that a rational assessment of ethicality should not depend on the identifiability of the victim of wrongdoing or the actual harm caused. Yet in four laboratory studies, we show that these factors have a systematic effect on how people judge the ethicality of the perpetrator of an unethical action. Specifically, we find that identifiability of the victim of wrongdoing and information about the outcome of wrongdoing influence both ethical judgments and decisions to punish wrongdoers. Our studies show that people judge behavior as more unethical when (1) identifiable versus statistical victims are involved and (2) the behavior leads to a negative rather than a positive outcome. We also find that people's willingness to punish wrongdoers is consistent with their judgments, and we offer preliminary evidence on how to reduce these biases.

*Keywords:* ethics; identifiability; judgment; outcome bias; unethical behavior

## Nameless + Harmless = Blameless:

## When Seemingly Irrelevant Factors Influence Judgment of (Un)ethical Behavior

In the unforgiving college admissions race, few people pause to question the common practice of giving preference to the close kin of alumni. This practice of “legacy admissions” has a real impact on college applicants. When a legacy is moved from below the admissions cutoff to above the cutoff, some other unidentified victim is probably transferred in the opposite direction. However, the fact that such victims are difficult to identify dampens any outrage that might be associated with such discriminatory practices. Now imagine that an applicant finds out that a legacy admit narrowly prevented her from being admitted to a certain college, but that she, the victim, was accepted by a school she preferred more. Again, few would be outraged by the rejection. Now imagine that the victim knew that a legacy admit kept her from being admitted to her first-choice school. In this case, the legacy admissions process would likely seem outrageous to outsiders. Yet we argue that corruption of an unbiased admissions process occurred equally in each of these cases.

Why has there not been greater uproar in the United States over the discriminatory practice of legacy admits at most universities? We believe that outrage is minimal due to the lack of identifiability of the victims of legacy admissions and lack of clarity concerning the outcomes of this unethical behavior. In this paper, we examine the effects of these two factors—the identifiability of the victim of wrongdoing and the outcomes of wrongdoing—on the harshness of people’s judgments of others’ unethical behavior. In the legacy admission context and beyond, we expect decision-makers to overlook others’ unethical behavior when such behavior hurts unknown victims and/or when it does not lead to negative consequences.

Thus, our research examines the effects of the identifiability of victims (Small & Loewenstein, 2003; Small & Loewenstein, 2005; Kogut & Ritov, 2005a; Kogut & Ritov, 2005b) and the effects of outcome information (Baron & Hershey, 1988) in ethical contexts.

### The Outcome Bias in the Realm of Ethics

Do people perceive instances of the same unethical behavior differently (and worthy of different punishment) if it has different outcomes? From the perspective of a rational decision-maker, when we are judging the quality of someone's decision, we should only evaluate that person's decision and decision-making process and should not evaluate a randomly determined outcome (Bazerman & Moore, 2008; Hastie & Dawes, 2001). However, previous research has shown that people exhibit an outcome bias: they judge the quality of decisions and the competence of decision-makers based on the outcomes of their decisions (Baron & Hershey, 1988; Allison, Mackie, & Messick, 1996; Marshall & Mowen, 1993; Mazzocco, Alicke, & Davis, 2004). This robust bias has been demonstrated in the realms of medicine, management, and corporate decision-making.

Investigating the influence of outcome information on ethical judgment, Gino, Moore, and Bazerman (2008) found that the same ethically questionable behaviors are judged as more unethical when they result in a negative outcome rather than a positive one. Using hypothetical scenarios, Gino et al. (2008) showed that even those who previously judged and rated the ethicality of a person's behavior changed their opinions after learning the outcome of the described behavior. As a result of these differences in judgment, we can expect people to judge an identical behavior differently depending on its outcome. In particular, we hypothesize that:

Hypothesis 1: People evaluate others' questionable behavior as more unethical when it leads to negative consequences than when it leads to positive consequences.

As philosophers and psychologists have suggested (e.g., North, 1987; Goldberg, Lerner, & Tetlock, 1999), people commonly use the judgments they form about the ethicality of an actor's behavior as inputs when determining whether the wrongdoer deserves punishment. Similarly, consumer behavior research has demonstrated that consumers' perceptions of a company's ethical behavior have a significant effect on their willingness to reward the firm by purchasing its products (e.g., Creyer & Ross, 1997). Consistent with this stream of research, Jones (1991) proposed an ethical decision-making model in which decision makers systematically evaluate the "intensity" of the essential features of the moral issue at hand when facing decisions in the ethical realm. The perceived overall intensity of a moral issue then determines the decision maker's judgment, intent, and subsequent behavior (Jones, 1991). Building on this prior work, we suggest that an observer's perception of the ethicality of a target's behavior will determine whether the target deserves to be punished for his or her behavior. Thus, we hypothesize that:

Hypothesis 2: People will be more likely to punish others' questionable behavior when it leads to negative consequences than when it leads to positive consequences.

Thus, the present research examines the effects of the outcome bias on both ethical judgments and decisions to punish others for their unethical acts. In addition, we investigate the effects of identifiability of the victim of wrongdoing on both judgment and behavior toward those behaving unethically.

### The Identifiable Victim Effect

The “identifiable victim effect” refers to the tendency of people to be far more concerned about and show more sympathy toward identifiable victims than statistical victims. Simply telling people that a specific victim exists increases caring, even when no personalizing information about the victim is available (Small & Loewenstein, 2003). Loewenstein, Small, and Strnad (2006) offer two types of reasons to explain why people show more concern for identified victims than for statistical victims (Kogut & Ritov, 2005a; Jenni & Loewenstein, 1997): affect-based and cognitive-based reasons.

On the affective level, identification decreases the social distance between victim and responder (Small & Loewenstein, 2005). Specifically, the same situation triggers greater sympathy when it involves just one identified victim than when it involves many non-identifiable victims – a singularity effect. Identification with a singular victim thus changes the affective reactions responders show to victims (Kogut & Ritov, 2005b). On the cognitive level, Friedrich, Barnes, Chapin, Dawson, Garst, and Kerr (1999) also emphasize the singularity of victims to explain the identifiable victim effect. They use the term “psychophysical numbing” to refer to the tendency for people to value lives less as the number of lives at risk increases. Because of this “drop in the bucket” effect, the authors argue that identifying a singular victim leads people to believe that the single life affected represents a disproportionate percentage of the total threat or problem. Independent of the increased sympathy and decreased psychological distance that identification may evoke, this framing argument implies that one victim creates a more salient cognitive reference frame than many individuals.

Extending the stream of research, Small and Loewenstein (2005) examined situations in which people feel they have been treated unfairly. Defining wrongdoing as “behaving in a self-interested way” instead of being cooperative, the researchers manipulated whether the wrongdoer is identified or unidentified. Small and Loewenstein (2005) show that people are more punitive toward identified wrongdoers than toward equivalent, but unidentified, wrongdoers. They also show that contributors react with greater anger toward identified non-contributors than toward unidentified non-contributors. In comparison, the present research investigates situations in which people observe unethical behavior that directly affects others but not themselves and manipulates whether the *victim* of wrongdoing, rather than the perpetrator of unethicity, is identified or unidentified. We predict that:

Hypothesis 3: People will judge greater unethicity when victims of wrongdoing are identified than when they are not.

Hypothesis 4: People will be more likely to punish others’ questionable behavior when victims of wrongdoing are identified than when they are not.

#### Reducing Biases in Ethical Judgment and Behavior

We are also interested in identifying strategies to reduce the judgment biases predicted in Hypotheses 1 through 4. We focus on the influence on judgment of making the issues of identifiability and outcomes transparent through joint evaluation. As compared to separate evaluation, which refers to settings in which only one alternative is considered, joint evaluation refers to contexts in which multiple examples or alternatives are considered at the same time (Bazerman, Loewenstein, & White, 1992). Joint evaluation allows decision-makers to assess various attributes simultaneously and gives them more information about the decision at hand than does separate evaluation. As a result, joint evaluation leads to more

reflective, reason-based choice (Hsee, Blount, Loewenstein, & Bazerman, 1999; Bazerman, Tenbrunsel, & Wade-Benzoni, 1998) than separate evaluation, which is more likely to be influenced by affective factors and to result in more intuitive judgments (Bazerman et al., 1998).

In our setting, we predict that effects of the identifiability of the victim of wrongdoing are present under conditions favoring intuition (separate evaluation). However, if the comparison between identifiable and unidentifiable victims of wrongdoing is made transparent under conditions favoring direct comparison and more reflective judgment (joint evaluation), the perception of the unethicity of the behavior causing harm to an identified victim will be reduced. Specifically, we hypothesize that:

Hypothesis 5: Joint evaluation will reduce the effect of the identifiability of the victim on ethical judgment.

Similarly, we predict a similar effect of joint evaluation in reducing the impact of outcome information on ethical judgment. Specifically, we expect the effects of the outcome bias to be present under conditions favoring intuition (separate evaluation). However, under conditions favoring direct comparison and more reflective judgment (joint evaluation), the effects of the outcome bias on ethical judgments will be reduced. We thus hypothesize that:

Hypothesis 6: Joint evaluation will reduce the effect of outcome information on ethical judgment.

### Overview of the Present Research

Four studies were conducted to test our hypotheses. Study 1 uses hypothetical scenarios and investigates the effects of identifiability of the victim of wrongdoing and the valence of outcome bias on ethical judgment, thus testing Hypotheses 1 and 3. Study 2

introduces a behavioral measure by examining whether the same factors influence decisions to punish those doing wrong, thus testing Hypotheses 2 and 4. Finally, Studies 3 and 4 use hypothetical scenarios and manipulate whether others' behavior is evaluated separately or jointly, thus testing Hypotheses 5 and 6.

### Study 1

Study 1 tests the hypotheses that people judge greater unethicity when victims are identifiable rather than statistical (i.e., main effect of identifiability of the victim of wrongdoing) and that people judge greater unethicity when a negative (rather than a positive) outcome results from the same unethical behavior (main effect of outcome information).

#### Methods

Participants. Two-hundred-three individuals (116 male, 87 female) participated in exchange for \$5. Most participants (93%) were students from local universities in Pittsburgh. The average age of participants was 23 ( $SD = 4.31$ ).

Procedure. The study employed a 2 (outcome information: positive vs. negative) X 2 (victim: unidentified vs. identified) design. The study consisted of a survey participants were asked to complete. The survey included two scenarios to evaluate and varied across the four experimental conditions. At the beginning of the study, participants were randomly assigned to one of the four experimental conditions. Participants in each condition were asked to read two scenarios. One scenario described the behavior of a doctor, and the other described the behavior of a real estate agent. The order in which scenarios were presented to participants was counterbalanced. Because all participants read both scenarios, the two scenarios were used as a repeated measure in the analyses presented below.

Manipulations. All four conditions included descriptions of the same elements of an arguably unethical action. Thus, objectively, the ethicality of the action was held constant across conditions. The *outcome information* was varied in the last line of the scenarios (as reported in brackets below). *Identifiability* was varied simply by including or not including a gender-neutral name for the victim of the action. In the unidentifiable condition, the victim was simply referred to as “the patient” (in the scenario describing the behavior of a doctor), or “the person” (in the scenario describing the behavior of a real estate agent). In the identified name condition, the victim was referred to using a gender-neutral name (i.e., Sam, Chris).

The scenario describing the behavior of a doctor read:<sup>1</sup>

(A person/Chris) has a 2-week history of low-back pain with worrisome symptoms. Although the pain has improved somewhat with bed rest, ibuprofen, and a heating pad at night, (this person/Chris) is concerned because such back pain has been persistent. (This person/Chris) is otherwise healthy and has no bone tenderness. Preoccupied, (this person/Chris) wants to see an orthopedic surgeon or get a magnetic resonance imaging scan. While practice guidelines in this case recommend referral to a specialist or imaging studies because of the likelihood of serious disease, the physician (this person/Chris) contacted only recommended some rest, an easier and less costly solution for the physician compared to the alternatives. The physician was indeed interested in saving money and time for himself. Doctors are compensated such that they collect a monthly fee for each patient and pay the costs of specialists out of their fee. (The patient’s/Chris’s) symptoms improve over time and after only a couple of weeks, [the pain completely disappears / the patient is in need of back surgery, which will very likely result in long-term effects such as pain and loss of mobility].

The other scenario described the behavior of a real estate agent:

(A person/Sam) has been hired for a new job and needs to move to a new city. (This person/Sam) has been looking at houses over the last two weeks and decided to make an offer on a house selling for \$350,000. Concerned with potential problems this house might have, (this person/Sam) wants to have the house inspected before buying it. During the house search, (this person/Sam) received recommendations from a real estate agent who works on a commission of 6% paid by the seller of the property. While practice guidelines in this case recommend referral to a list of three thorough non-biased inspectors, the real estate agent only recommended one home inspector, who the agent knows quite well. This creates a potential conflict of interest as the

---

<sup>1</sup> In the scenarios, victim identification appears in parentheses ( ) and outcome information appears in brackets [].

appraiser knows that the real estate agent would be disappointed if the appraiser found significant problems that endangered the transaction. And the agent is expecting a \$21,000 commission. Indeed, the real estate agent views a thorough and non-biased home inspection as a threat to the sales commission. The appraiser finds no significant problem. (The person/Sam) ends up buying the house for \$350,000, and [lives there happily for five years with no house problem / experiences serious problems with flooding in the basement soon after. (The person/Sam) invests \$20,000 to finish the basement, which is flooded and destroyed a few months later.]

After reading each scenario, participants were asked to answer three questions. The first question asked about the extent to which participants found the behavior of the doctor or the real-estate agent, as described in the scenario, to be unethical. The second question asked participants to indicate the extent to which they found the behavior of the doctor or the real-estate agent, as described in the scenario, to be intentional. Finally, the last question asked participants to indicate how much blame they felt toward the wrongdoer. Participants answered each question by using a 7-point scale (first question: 1 = very ethical, 7 = very unethical; second question: 1 = not intentional at all, 7 = very intentional; third question: 1 = not at all, 7 = very much).

## Results

Correlations. A partial correlation analysis (controlling for both outcome information and identifiability of the victim) revealed a significant and positive correlation between ratings of unethicity and intentionality of the described behavior ( $r = .53, p < .001$ ). Ratings of unethicity were also positively correlated with ratings of blame ( $r = .53, p < .001$ ).

Effects of outcome information and identifiability of victims. Hypothesis 1 and Hypothesis 3 predict that people's judgments of unethicity will vary based on the type of outcome observed, as well as the identifiability of the victim, respectively. We tested these hypotheses by comparing participants' ratings for unethicity across conditions.

Unethicality. We used the unethicality ratings as the dependent variable in an analysis of variance (ANOVA) in which outcome information (positive vs. negative) and identifiability of the victim (unidentified vs. identified) served as between-subjects factors (repeated-measure on scenario). This analysis revealed a significant main effect for outcome information,  $F(1, 198) = 119, p < .001, \eta^2 = .38$ : the rating for unethicality was higher in the negative-outcome ( $M = 6.20, SD = 0.98$ ) than in the positive-outcome condition ( $M = 4.46, SD = 1.27$ ). The main effect for identifiability of the victim was also significant,  $F(1, 198) = 5.34, p < .05, \eta^2 = .03$ : the rating for unethicality was lower in the unidentified victim condition ( $M = 5.14, SD = 1.46$ ) than in the identified victim condition ( $M = 5.54, SD = 1.37$ ). The interaction was not significant ( $p = .41$ ). Additional analyses revealed that in the negative-outcome condition, the rating for unethicality was higher for identified victims ( $M = 6.45, SD = 0.66$ ) than for unidentified victims ( $M = 5.95, SD = 1.16; t(100) = 2.65, p = .009$ ). In the positive-outcome condition, by contrast, there was no significant difference in the unethicality ratings between identified ( $M = 4.57, SD = 1.27$ ) and unidentified victims ( $M = 4.35, SD = 1.28; t(99) < 1, p = .38$ ).

Intentionality. We used the intentionality ratings as the dependent variable in an ANOVA in which outcome information and identifiability of the victim served as between-subjects factors (repeated-measure on scenario). This analysis revealed a significant main effect for outcome information,  $F(1, 198) = 15.93, p < .001, \eta^2 = .07$ : the rating for intentionality was higher in the negative-outcome ( $M = 5.87, SD = 1.07$ ) than in the positive-outcome condition ( $M = 5.15, SD = 1.44$ ). The main effect for identifiability of the victim was also significant,  $F(1, 198) = 11.84, p = .001, \eta^2 = .06$ : the rating for intentionality was

lower in the unidentified victim outcome ( $M = 5.22$ ,  $SD = 1.50$ ) than in the identified victim condition ( $M = 5.83$ ,  $SD = 0.99$ ). Note that the interaction term was insignificant ( $p = .79$ ).

Blame. We conducted a similar 2 (outcome information) X 2 (identifiability of the victim) ANOVA using blame ratings as the dependent variable (repeated-measure on scenario). This analysis revealed a significant main effect for outcome information,  $F(1, 198) = 155$ ,  $p < .001$ ,  $\eta^2 = .44$ : the rating for blame was higher in the negative-outcome ( $M = 5.87$ ,  $SD = 1.01$ ) than in the positive-outcome condition ( $M = 3.76$ ,  $SD = 1.36$ ). The main effect for identifiability of the victim was also significant,  $F(1, 198) = 4.17$ ,  $p < .05$ ,  $\eta^2 = .02$ : the rating for blame was lower in the unidentified victim outcome ( $M = 4.64$ ,  $SD = 1.60$ ) than in the identified victim condition ( $M = 5.02$ ,  $SD = 1.53$ ). The interaction term was not significant ( $p = .57$ ).

## Discussion

The results of our first study provide support for Hypotheses 1 and 3. People evaluated others' questionable behavior as more unethical when it led to negative consequences and when victims of wrongdoing were identified.

### Study 2

Study 2 examined the influence of outcome information and identifiability of the victim of wrongdoing on individuals' behaviors toward the perpetrator of ethically questionable actions. In this study, participants observed the behavior of another participant and had the opportunity to punish this participant by reducing his/her payoff.

## Methods

Participants. Two-hundred twenty-three individuals (127 male, 96 female) participated in the study for monetary compensation. The average age of participants was 23 ( $SD = 2.93$ ). Most participants (79%) were students from local universities in Pittsburgh.

Procedure. Groups of between 3-6 people participated in the study which was conducted on computers. The experiment was described to participants as a study of money allocation decisions. They were told that they would be randomly assigned to a role and that, depending on their role, they would decide how to allocate money to themselves or to other participants in a different role. Participants were assigned one of six different roles: Player A, Player B, and one of four types of Player C. In the sessions involving three participants, participants were randomly assigned to one of three roles: Player A, Player B, and one type of Player C. When more than three participants were part of the session, they were randomly assigned to the role of Player A, Player B, and different types of Player C.

Participants in the role of Player A chose payments for themselves and another participant (Player B) in the same study. Specifically, Player As had to choose between the following two options: “1) You receive \$5, and Player B receives \$5; 2) You receive \$6, and the experimenter will toss a fair coin. If heads, then Player B will receive \$5. If tails, then the other participant will receive nothing.” Participants in the role of Player B were truthfully told that their payment would depend upon the choice of Player A. Participants in the role of Player A and Player B were matched so that Player A’s decisions did determine Player B’s payoff in the study.

For participants playing the role of Player C, we manipulated both identifiability and outcome information. The instructions on the computer screen informed participants of their role (Player C) and told them that they would receive \$5 for their participation. Identifiability

was manipulated solely by referring to Player B with a gender-neutral name (Chris) in the identifiable condition. Player Cs received the following instructions for the game: “In this study, you will observe the actions of two randomly chosen participants in the same study – Player A and [Player B/Chris]. Player A will be asked to consider the following two options: Option 1) A receives \$5, and [Player B/Chris] receives \$5. Option 2) A receives \$6, and the experimenter will toss a fair coin. If heads, then [Player B/Chris] will receive \$5. If tails, then [Player B/Chris] will receive nothing.” Player Cs then waited for Player A and Player B to make their decisions. This waiting time was actually fictitious since the choice of Player A that Player Cs saw on their computer screen was determined based on their type (remember that there were four types of possible roles for Player C depending on the experimental condition). We used this fictitious waiting time to lead Player Cs to believe they were actually playing with other participants in the study assigned to the role of Player A and Player B. In reality, they were not. Indeed, Player Cs were always told that Player A chose the second option, namely the one in which Player A receives \$6.

Outcome information was manipulated by changing the information participants in the role of Player C received about the coin toss. In the negative-outcome information condition, participants were told, “Player A has chosen Option 2. The experimenter has tossed the coin, and the result of the coin toss was tails. Player A will thus receive \$6, and [Player B/Chris] will receive no payment for this study.” In the positive-outcome information condition, participants instead were told, “Player A has chosen Option 1. The experimenter has tossed the coin, and the result of the coin toss was heads. Player A will thus receive \$6, and [Player B/Chris] will also receive \$5 for this study.”

In all conditions, after receiving this information, participants in the role of Player C were given the option to punish Player A. However, punishing Player A created costs for Player C. This task is an adaptation of the punishment game created by Eckel and Grossman (1996), which in turn was a restricted version of a dictator game (Roth, 1995). A task similar to our game was used by Dana, Weber and Kuang (2008). In our game, Player Cs were faced with the following choices:

- Pay \$0.00, and thus reduce Player A's payoff by \$0.00
- Pay \$0.05, and thus reduce Player A's payoff by \$0.25
- Pay \$0.10, and thus reduce Player A's payoff by \$0.50
- Pay \$0.15, and thus reduce Player A's payoff by \$0.75
- Pay \$0.20, and thus reduce Player A's payoff by \$1.00
- Pay \$0.25, and thus reduce Player A's payoff by \$1.25

Participants in the role of Player C were told that the money they decided to pay to reduce Player A's payoff would be deducted from their \$5 participation fee at the end of the study. Participants were paid based on their decisions. After the Money Allocation Study was over, participants were informed that they would also be paid \$2 for filling out a short, unrelated questionnaire. We included this final questionnaire to make sure each participant received at least \$2 for their participation; depending on the decisions of participants in the role of Player A, participants in the role of Player B could have ended up with nothing. While the study only lasted about 10 minutes, we wanted everybody to leave the lab with a positive payoff.

While we were only interested in the decisions of participants in the role of Player C, we included Player A and Player B so that the game would be perceived as realistic.

## Results

For our analyses, we were interested in the decisions made by participants in the role of Player C ( $N = 145$ ). We first examined their punishment decisions. Figure 1 reports the number of responses for each level of punishment by condition.

-----  
 Insert Figure 1 about here  
 -----

We used the punishment level as the dependent variable in a 2 (outcome information) X 2 (identifiability) ANOVA. To punish Player A, participants were willing to reduce their own payoff by a higher amount in the negative-outcome ( $M = 0.11$ ,  $SD = 0.11$ ) than in the positive-outcome condition ( $M = 0.03$ ,  $SD = 0.06$ ),  $F(1, 141) = 36.81$ ,  $p < .001$ ,  $\eta^2 = .21$ . This result supports Hypothesis 2, which predicted that people will be more likely to punish others' questionable behavior when it leads to negative consequences than when it leads to positive consequences. The main effect for identifiability of the victim was also significant,  $F(1, 141) = 8.99$ ,  $p < .01$ ,  $\eta^2 = .06$ : the punishment level was lower in the unidentified victim-outcome ( $M = 0.05$ ,  $SD = 0.08$ ) than in the identified condition ( $M = 0.09$ ,  $SD = 0.11$ ). This result supports Hypothesis 4, which predicted that the same questionable behavior will lead to higher levels of punishment when the victim of wrongdoing is identified than when he/she is not. Note that the outcome information and identifiability interaction was not significant,  $F(1, 141) = 1.52$ ,  $p = .22$ ,  $\eta^2 = .01$ .

We examined whether outcome information and identifiability of the victim influence the decision to punish or not, without regard for the amount of the punishment. We created a dummy variable that was equal to 1 if participants in the role of Player C decided to punish Player A by reducing their own payoff and 0 if they decided not to punish Player A. The

number of responses for each of these two sets of choices by condition is reported in Table 1. The valence of the outcome information influenced the percentage of responses in favor of the choice of punishing versus not punishing,  $\chi^2 (N = 145) = 26.56, p < .001$ . Identifiability also influenced the percentage of responses in favor of punishing versus not punishing,  $\chi^2 (N = 145) = 3.88, p < .05$ .

-----  
 Insert Table 1 about here  
 -----

### Discussion

Our second study provides further support for the effects of identifiability of the victim of wrongdoing and the nature of outcome information in the ethical realm by showing that these factors influence not only ethical judgment (as in Study 1), but also the decision to punish others who do wrong. Specifically, the result of Study 2 show that people are more likely to punish others for their unfair behavior toward a third party when the unfair behavior produced negative rather than positive consequences and when the third party being harmed is identified rather than unidentified. These finding are particularly interesting in light of the fact that punishing others in this study meant not only reducing the payoff of the perpetrator of wrongdoing but also reducing one's own payoff.

### Study 3

Study 3 examined ways to reduce the effects of identifiability of the victim of wrongdoing demonstrated in our first two studies. Specifically, we looked at the influence of identifiability of the victim of wrongdoing under two distinct conditions: separate evaluation

and joint evaluation. While the former condition favors intuitive judgment, the latter favors more reflective, analytical judgment.

We used the joint/separate manipulation of Bazerman, Loewenstein, and White (1992) to investigate the effects of identifiability of the victim of wrongdoing in a setting in which the outcome of the wrongdoer's decisions was negative. We predicted that effects of identifiability of the victim of wrongdoing would be present under conditions favoring intuition (separate evaluation). However, if the comparison between identifiable and unidentifiable victims of wrongdoing is made transparent under conditions favoring direct comparison and more reflective judgment (joint evaluation), the effect of identifiability should be reduced.

### Methods

We recruited participants from various locations in the city of Pittsburgh to complete one of three versions of a survey in exchange for a candy bar. A total of 150 participants, evenly divided across the three conditions, completed the survey. All participants read about scenarios with a negative outcome. Participants were randomly assigned to one of three conditions: *separate evaluation – negative outcome/identified victim* (SE-NI), *separate evaluation – negative outcome/unidentified victim* (SE-NU), or *joint evaluation* (JE). In the SE-NU condition, participants read the following scenario; in the SE-NI condition, the text in brackets replaced the phrase “a/this person”:

A person [Sam] has a 2-week history of low-back pain with worrisome symptoms. Although the pain has improved somewhat with bed rest, ibuprofen, and a heating pad at night, this person [Sam] is concerned because such back pain has been persistent. This person [Sam] is otherwise healthy and has no bone tenderness. Preoccupied, this person [Sam] wants to see an orthopedic surgeon or get a magnetic resonance imaging scan. While practice guidelines in this case recommend referral to a specialist or imaging studies because of the likelihood of serious disease, the physician this person contacted only recommended some rest, an easier and less costly solution for the

physician compared to the alternatives. The physician was indeed interested in saving money and time for himself. Doctors are compensated such that they collect a monthly fee for each patient and pay the costs of specialists out of their fee. The patient's [Sam's] symptoms worsen over time and after only a couple of weeks, the patient [Sam] is in need of back surgery, which will very likely result in long-term effects such as pain and loss of mobility.

All participants in the SE-NI and SE-NU conditions were then asked to rate how unethical the actions of the described doctor were on a 9-point scale anchored by *Not at all Unethical* (-4) and *Very Unethical* (+4). In the JE condition, participants read both the scenario used in the SE-NI condition and the scenario used in the SE-NU condition. We counterbalanced the order in which the scenarios were presented to the participants in the JE condition. After reading both scenarios, participants in the JE condition were asked to rate how unethical the actions of the described doctor were on a 9-point scale anchored by *Not at all Unethical* (-4) and *Very Unethical* (+4).

### Results

We first compared the ratings of unethicality between the two separate evaluation conditions (SE-NI and SE-NU). Participants rated the behavior of the doctor as significantly more unethical when the victim of wrongdoing was identified ( $M = 2.98$ ,  $SD = 1.13$ ) than when the victim was unidentified ( $M = 1.38$ ,  $SD = 2.32$ ),  $t(98) = 4.38$ ,  $p < .001$ .

Next, we compared the ratings of unethicality for the behavior described in the two scenarios participants received in the joint evaluation condition. These participants rated the behavior of the doctor as more unethical when the victim of wrongdoing was identified ( $M = 2.28$ ,  $SD = 1.98$ ) than when the victim was unidentified ( $M = 1.98$ ,  $SD = 2.39$ ),  $F(1, 49) = 6.04$ ,  $p = .02$ ,  $\eta^2 = .11$ .

When the victim of wrongdoing was identified, ratings of unethicality dropped significantly, from 2.98 in the separate evaluation condition to 2.28 in the joint evaluation

condition,  $t(98) = 2.17, p < .05$ . By contrast, when the victim of wrongdoing was unidentified, ratings of unethicity did not significantly differ between the separate evaluation and the joint evaluation conditions ( $t[98] = 1.28, p = .21$ ). Essentially, the joint evaluation condition led participants to converge the ethicality judgments across the two identifiability conditions. Taken together, these results provide support for Hypothesis 5, which predicted that joint evaluation would reduce the effect of identifiability of the victim on ethical judgment.

We also computed the difference in unethicity scores between the identified victim and unidentified victim scenarios in both the joint and separate evaluation conditions.<sup>2</sup> The difference score was significantly lower in the joint evaluation condition than in the separate evaluation condition (0.30 vs. 1.60,  $t [98] = -3.15, p = .002$ ). Note also that 82% of the participants in the joint evaluation condition reported the same ratings for the unethicity of the described behavior in the identified victim and unidentified victim conditions.

### Discussion

The results of our third study support our hypothesis that, when evaluated separately, unethical behavior that harms an identified person would be considered more unethical than unethical behavior that harms an unidentified person. Yet, as our results show, when these two behaviors are evaluated jointly, the bias is reduced. These results suggest that the impact of identifiability can be reduced by using joint evaluation to make the transparency of the identification issue clear.

---

<sup>2</sup> In the case of the two separate evaluation conditions, we used a random matching procedure to match the answers to the two scenarios of SE-NI and SE-NU. We repeated this analysis 20 times with a different match each time. The nature and significance of the results did not change.

### Study 4

Study 4 investigated the influence of the valence of the outcome bias when the victim of wrongdoing is identified under two distinct conditions, namely separate evaluation and joint evaluation. Similar to our predictions for Study 3, in Study 4 we expect the effects of the outcome bias to be present under conditions favoring intuition (separate evaluation). However, under conditions favoring direct comparison and more reflective judgment (joint evaluation), we expect the effects of the outcome bias on ethical judgments to be reduced.

#### Methods

As in Study 3, we recruited participants from various locations in the city of Pittsburgh to complete one of three versions of a survey in exchange for a candy bar. A total of 148 participants completed the survey. Participants were randomly assigned to one of three conditions: *separate evaluation – negative outcome/identified victim* (SE-NI), *separate evaluation – positive outcome/identified victim* (SE-PI), or *joint evaluation* (JE). In the SE-NI condition, participants read the following scenario:

Sam has a 2-week history of low-back pain with worrisome symptoms. Although the pain has improved somewhat with bed rest, ibuprofen, and a heating pad at night, Sam is concerned because such back pain has been persistent. Sam is otherwise healthy and has no bone tenderness. Preoccupied, Sam wants to see an orthopedic surgeon or get a magnetic resonance imaging scan. While practice guidelines in this case recommend referral to a specialist or imaging studies because of the likelihood of serious disease, the physician Sam contacted only recommended some rest, an easier and less costly solution for the physician compared to the alternatives. The physician was indeed interested in saving money and time for himself. Doctors are compensated such that they collect a monthly fee for each patient and pay the costs of specialists out of their fee. *Sam's symptoms worsen over time and after only a couple of weeks, Sam is in need of back surgery, which will very likely result in long-term effects such as pain and loss of mobility.* [Sam's symptoms improve over time and after only a couple of weeks, the pain completely disappears.]

In the SE-PI condition, the last sentence (reported here in italics for clarification purposes) was substituted with the text in brackets.

All participants were then asked to rate how unethical the actions of the described doctor were on a 9-point scale anchored by *Not at all Unethical* (-4) and *Very Unethical* (+4). In the JE condition, participants read both the scenario used in the SE-NI condition and the scenario used in the SE-PI condition. As in Study 3, we counterbalanced the order in which the scenarios were presented to participants in the joint-evaluation condition. After reading both scenarios, participants were asked to rate how unethical the actions of the described doctor were on a 9-point scale anchored by *Not at all Unethical* (-4) and *Very Unethical* (+4).

### Results

We first compared the ratings of unethicality in the two separate evaluation conditions. Respondents rated the behavior of the doctor toward an identified victim as significantly more unethical when it resulted in a negative outcome ( $M = 3.08$ ,  $SD = 0.95$ ) than when it resulted in a positive outcome ( $M = 0.76$ ,  $SD = 2.45$ ),  $t(97) = 6.20$ ,  $p < .001$ .

Next, we compared ratings of unethicality for the behavior described in the two scenarios that participants received in the joint evaluation condition. These participants rated the behavior of the doctor as more unethical when it resulted in a negative outcome ( $M = 1.93$ ,  $SD = 2.24$ ) than when it resulted in a positive outcome ( $M = 1.27$ ,  $SD = 2.38$ ),  $F(1, 44) = 6.67$ ,  $p = .01$ ,  $\eta^2 = .13$ .

When the outcome of the doctor's behavior was negative, ratings of unethicality dropped significantly, from 3.08 in the separate evaluation condition to 1.93 in the joint evaluation condition,  $t(94) = 3.25$ ,  $p = .002$ . By contrast, when the outcome of the doctor's behavior was positive, ratings of unethicality did not significantly differ between the separate evaluation and the joint evaluation conditions ( $t[95] < 1$ ,  $p = .36$ ). Again, the joint evaluation condition led to less bias, this time converging the ethicality judgments across the two

outcome conditions. Taken together, these results support Hypothesis 6, which predicted that joint evaluation would reduce the effect of outcome information on ethical judgments.

As in Study 3, we conducted additional analyses to further investigate the influence of the separate and joint evaluation of a wrongdoer's unethicity in the case of a negative versus positive outcome for the wrongdoer's behavior. We computed the difference in unethicity scores between the negative and positive outcome scenarios in both the joint and the separate evaluation.<sup>3</sup> The difference score was significantly lower in the joint evaluation condition than in the separate evaluation condition ( $t [96] = -3.63, p < .001$ ). Note also that 65% of the participants in the joint evaluation condition reported the same ratings for the unethicity of the described behavior in the negative and positive outcome conditions.

## Discussion

The results of Study 4 provide evidence for the robustness of the outcome bias in ethical realms. Consistent with our predictions, the valence of outcome information influences ethical judgments when separate evaluation is used. This effect is reduced, but not eliminated, in the case of joint evaluation.

### General Discussion

Our research systematically investigated how irrelevant factors that should not enter into an individual's decision-making process nonetheless have an impact on his or her judgments and behavior in the ethical realm. In the same vein, recent studies by Gino et al. (2008) demonstrated that people tend to base their ethical judgments of others' behavior on the type of consequences that results. The current research contributes to this stream of work by showing that outcome information influences not only one's judgments of others'

---

<sup>3</sup> As in Study 3, also in this case, we used a random matching procedure to match the answers to the two scenarios of SE-NI and SE-PI. We repeated this analysis 20 times with a different match each time. The nature and significance of the results did not change.

ethicality, but also one's own decision to punish others for their unethical behavior, even if this means reducing one's own payoff.

Our work also contributes to the stream of research on identifiable victims. This body of work has shown that people tend to be far more concerned about and show more sympathy toward identifiable victims than statistical victims. Our research builds on these findings to show that people tend to consider others' questionable behaviors to be more unethical when the victim of wrongdoing is identified versus unidentified and that people also punish wrongdoers more harshly in the former case than in the latter. Our studies, we believe, also make a broader contribution to research on ethical judgment and decision making. In addition to providing evidence for both the outcome bias and identifiability of the victim effects in the ethical realm, we demonstrate that these biases can be reduced when the judgment process is approached using joint evaluation.

We see several directions for future research building on the present work. First, further work could investigate other strategies or tools that decision makers could use to reduce and possibly even eliminate the biases demonstrated here. For instance, future studies could include conditions in which the description and evaluation phases happen sequentially in the case of joint evaluation. That is, both outcomes (and the behaviors that lead to them) first would be described or presented to the decision maker, who would then evaluate them by comparing one to the other directly. Another potentially interesting direction for future research would be to investigate whether the same effects would hold in a setting (possibly a real-world setting) with repeated interactions, where judges who can punish others have ongoing relationships with wrongdoers or with the victims of wrongdoing.

The research presented here also has significant practical implications. The effects demonstrated in our studies might lead people to blame others too harshly for making sensible decisions that have unlucky outcomes or that harm a clearly identifiable victim. Our results could also explain the slow reactions people tend to have when observing others' unethical behavior. Too often, we let others' ethically questionable decisions slide over a period of time until they result in negative consequences, even when such consequences can be easily predicted (Bazerman & Watkins, 2004). Similarly, we tend to tolerate others' ethically questionable decisions until we have identified a clear victim. Thus, the present research has implications for both organizations and public institutions. As our studies show, decision-makers should anticipate being judged less for the ethics of their actions than for the consequences of those actions and the identifiability of the victim of their wrongdoing. No matter how ethical the decisions of a manager or a company may be, judges (such as customers, citizens, or employees) might punish that manager or company if things go wrong or if victims are clearly identified.

Rawls (1971) proposed that fairness should be assessed under a "veil of ignorance"—that is, we ideally should judge a situation without knowing the role we ourselves play in it. Extending this standard to our context, we should not only ignore the role we play in a situation, but should also ignore normatively irrelevant information concerning the identifiability of the victim of wrongdoing and the outcomes that result from ethically questionable practices. Yet, we have shown, empirically, that the effects of outcome and victim information are potentially large. Under U.S. law, the difference between attempted murder and first-degree murder leads to very different punishments—all hinging upon a difference of outcome even in the face of identical action and intention. In college admissions

offices, identifying the victims of admissions policies that favor certain candidates' applications over those of others may lead to more damage in the ethicality of the decision than retribution. Our claim is that the decision to withhold or disclose information about the victims and outcomes of a behavior can be a powerful determinant of the ethical perception of that behavior—and such a decision should not be overlooked as trivial.

### Conclusion

Most of us regularly make ethical judgments about others' behavior and make decisions regarding whether or not to punish others' unethical behavior. Although many of us know how we would *rationally* like to behave in these situations, little prior research has explored the systematic errors we commit in the process of evaluating others' unethical behavior and acting upon it. The present research focused on the effects of both the outcome valence of unethical acts and the identifiability of the victim of wrongdoing on ethical judgments and decisions to punish unethical behavior. By investigating these factors, we identify conditions under which our ethical judgments and behaviors are biased. In addition, our research considers ways to help decision makers reduce these biases and recognize that nameless plus harmless does not equal blameness.

## References

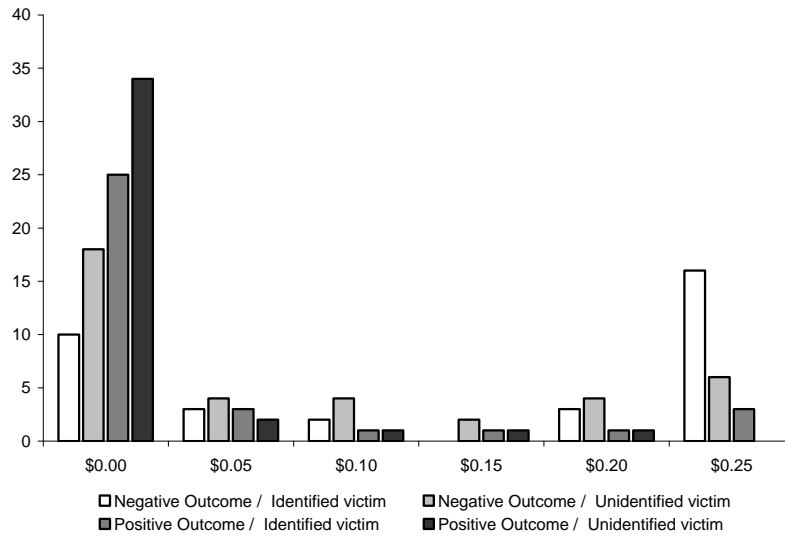
- Allison, S. T., Mackie, D. M., & Messick, D. M. (1996). Outcome biases in social perception: Implications for dispositional inference, attitude change, stereotyping, and social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 53-93). New York: Academic.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54, 569-579.
- Bazerman, M. H., Loewenstein, G., Blount White, S. (1992). Reversals of preference in allocation decisions: Judging an alternative versus choosing among alternatives. *Administrative Science Quarterly*, 37, 220-240.
- Bazerman, M. H., & Moore, D. A. (2008). *Judgment in managerial decision making* (7th ed.). New York: Wiley.
- Bazerman, M. H., Tenbrunsel, A. E., Wade-Benzoni, K. A. (1998). Negotiation with yourself and losing: Understanding and managing competing internal preferences. *Academy of Management Review*, 23, 225-241.
- Creyer, E. H., & Ross Jr., W.T. (1997). The influence of firm behavior on purchase intention: do consumers really care about business ethics? *Journal of Consumer Marketing*, 14, 421-432.
- Bazerman, M. H., & Watkins, M. D. (2004). *Predictable surprises. The disasters you should have seen coming and how to prevent them*. Boston: Harvard Business School Press.
- Dana, J., Weber, R. and J. X. Kuang. (2008). Exploiting Moral Wriggle Room: Behavior Inconsistent with a Preference for Fair Outcomes. In press, *Economic Theory*.

- Eckel, C. C., & Grossman, P. J. (1996). The relative price of fairness: gender differences in a punishment game. *Journal of Economic Behavior & Organisation*, 30, 143-158.
- Friedrich, J., Barnes, P., Chapin, K., Dawson, I., Garst, V., & Kerr, D. (1999). Psychophysical numbing: When lives are valued less as the lives at risk increase. *Journal of Consumer Psychology*, 8(3), 277-299.
- Gino, F., Moore, D. A., & Bazerman, M. H. (2008). No harm, no foul: The outcome bias in ethical judgments. *Under review*.
- Goldberg, J. H., Lerner, J. S., & Tetlock, P. E. (1999). Rage and reason: The psychology of the intuitive prosecutor. *European Journal of Social Psychology*, 29, 781-795.
- Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks, Calif.: Sage.
- Hsee, C. K., Blount, S., Loewenstein, G., Bazerman, M. (1999). Preference reversals between joint and separate evaluations of options a review and theoretical analysis. *Psychological Bulletin*, 125, 576-590.
- Jenni, K. E., & Loewenstein, G. F. (1997). Explaining the identifiable victim effect. *Journal of Risk and Uncertainty*, 14, 235-257.
- Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue contingent model. *Academy of Management Review*, 16 (April), 366-395.
- Kogut, T., & Ritov, I. (2005a). The “identified victim” effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, 18, 157-167.
- Kogut, T., & Ritov, I. (2005b). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, 97, 106-116.

- Loewenstein, G., Small, D. A., & Strnad, J. (2006). Statistical, identifiable, and iconic victims. In E. J. McCaffery & J. Slemrod (Eds.), *Behavioral public finance* (pp. 32–46). New York: Russell Sage Foundation Press.
- Marshall, G. W., & Mowen, J. C (1993). An experimental investigation of the outcome bias in salesperson performance evaluations. *Journal of Personal Selling and Sales Management*, 8, 31-47.
- 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.
- North, J. (1987). Wrongdoing and forgiveness. *Philosophy*, 62 499-508.
- Rawls, J. (1971). *A theory of justice*. Cambridge Mass.: Harvard University Press.
- Roth, A. E., (1995). Bargaining experiments, In: John Kagel and Alvin E. Roth, Eds., *Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.
- Small, D. A., & Loewenstein, G. (2003). Helping the victim or helping a victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, 26(1), 5–16.
- Small, D. A., & Loewenstein, G. (2005). The devil you know: The effect of identifiability on punitiveness. *Journal of Behavioral Decision Making*, 18(5), 311–318.

Figures

Figure 1. Number of responses on the decision to punish Player A by condition, Study 2.



## Tables

Table 1

*Number of responses for the choice of punishing or not punishing by condition, Study 2*

	Negative Outcome		Positive Outcome	
	Identifiable	Unidentifiable	Identifiable	Unidentifiable
Not punish	29%	47%	74%	87%
Punish	71%	53%	26%	13%