

This article was downloaded by: [Harvard College]

On: 12 March 2014, At: 10:54

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Ethics & Behavior

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/heb20>

Ethical Situations and Their Effects on Judgments of Punishment

Ethan Ludwin-Peery^a & Dustin Tingley^b

^a School of Cognitive Science, Hampshire College

^b Government Department, Harvard University

Accepted author version posted online: 25 Sep 2013. Published online: 10 Mar 2014.

To cite this article: Ethan Ludwin-Peery & Dustin Tingley (2014) Ethical Situations and Their Effects on Judgments of Punishment, *Ethics & Behavior*, 24:3, 253-262, DOI: [10.1080/10508422.2013.845534](https://doi.org/10.1080/10508422.2013.845534)

To link to this article: <http://dx.doi.org/10.1080/10508422.2013.845534>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Ethical Situations and Their Effects on Judgments of Punishment

Ethan Ludwin-Peery

*School of Cognitive Science
Hampshire College*

Dustin Tingley

*Government Department
Harvard University*

We compared the punitiveness of two groups following a manipulation in which participants were either able to cheat on a simple number-matching task, by taking more money than they rightfully earned, or prevented from doing so on the same task. After completing the task, participants read a number of small vignettes of politicians who had acted questionably, and then were asked to rate the scenarios on both how wrong the behavior was and how much punishment it deserved. Participants given the opportunity to cheat with impunity were significantly less punitive when judging questionable behavior on the part of elected officials.

Keywords: social norms, moral standards, unethical behavior, exposure, moral judgment, empathy, cognitive dissonance

INTRODUCTION

Although classical research considered ethical decision making to be a purely rational exercise in self-interest (Becker, 1968), new research is challenging this notion. The classical view is that decisions are made via a cost–benefit analysis based on estimations of three factors—payoff, probability of being caught, and the magnitude of punishment if caught. Recent research, however, suggests that the majority of individuals make ethical decisions without fully engaging in cost–benefit trade-off analysis. Furthermore, the context of the decision making leads to shifts in judgment even when the contextual factors do not affect any of the expected rational inputs (Mazar, Amir, & Ariely, 2008a).

A large body of recent work has focused on this problem of ethical decision making, especially operationalized as the decision to cheat. Through this work, two important discoveries have been made. First, the underpinnings of dishonest behavior are much more complex than previously believed—a number of subtle factors, especially concerning the ethical behavior of

Correspondence should be addressed to Ethan Ludwin-Peery, 2103 Notch Road, Jericho, VT 05465. E-mail: ethan@peery.com

others, can sway dishonest action considerably (Gino, Ayal, & Ariely, 2009). Second, given the right circumstances, cheating is something that nearly everyone does (Mazar et al., 2008a).

Perceptions of what is normal or acceptable ethical behavior, based on immediate experience, can significantly affect ethical decision making (Keizer, Lindenberg, & Steg, 2008). For this study, we examined if the experience of unethical behavior going unchecked changed subjects' perception of acceptable ethical behavior.

ETHICS AND SOCIAL NORMS

One of the most persistent factors affecting ethical decision making is our social structures and social norms. Many of the factors involved with ethical decision making appear to be cues to social interactions—darkness is considered a measure of anonymity (Zhong, Bohns, & Gino, 2010), and eyespots appear to invoke the sensation of being observed (Bateson, Nettle, & Roberts, 2006), prompting decreased and improved honesty, respectively. Social pressures, even when indirect, are also effective at modifying ethical behavior. In the context of a study in which participants are paid for correct performance, they tend to overreport their scores (effectively stealing from experimenters) when given the chance, and levels of cheating increase when an individual (actually a confederate) is seen cheating without consequence. Levels of cheating drop considerably, however, when the obvious cheater is from a social outgroup—in one case, wearing a T-shirt from a rival university (Gino et al., 2009).

Similar results are found when the research is extended to other forms of social awareness. After being cued/primed to expected rules of behavior—being asked to list the Ten Commandments, in one example—participants appeared to not cheat at all. Their performance was indistinguishable from the control condition, in which they were primed with the same reminder but were not given the opportunity to cheat. In addition to more universal cultural rules of behavior such as the Ten Commandments, standards of behavior for specific groups seem effective as well. Many universities, for example, have academic honor codes, an explicit set of rules and principles that define what constitutes acceptable academic behavior. The presence of an honor code reminder was also found to be effective at discouraging cheating behavior, even when the participating institutions did not actually have an honor code (Mazar et al., 2008a).

IDENTIFICATION AND EMPATHY

As just mentioned, ingroup/outgroup status can have a significant effect on how people treat one another. Ingroup individuals are treated with greater empathy than outgroup individuals, and individuals from the same group appear more inclined to help one another (Stürmer, Snyder, Kropp, & Siem, 2006; Tajfel, 1970). This effect appears even when the ingroup/outgroup distinction is trivial or even invented. Stürmer et al. (2006) told their participants that there were two types of cognitive strategies—"detailed" and "global." All participants were told that they were detailed perceivers, and went on to be more helpful to other supposed detailed perceivers, even though the distinction had just been learned and was, in fact, entirely spurious.

A 2000 study by Gault and Sabini found that trait empathy—empathy that is a *trait* of the person in question, identifying a characteristic of their personality, and not their emotional state at

the moment—caused participants to prefer a victim-support solution to a disaster over a punitive solution, suggesting that punitiveness may be reduced in general by empathy.

THE CURRENT STUDY

This study seeks to examine the relationship between social norms for ethical behavior and judgments of wrongness and punitiveness. Because of frequent problems with the ethical behavior of politicians and elected officials, a good deal of research has focused on ethical decision making in politics. We have chosen political punitiveness because we wished to continue investigating this important issue. We do, however, expect this measure to generalize well, and not simply apply to judgments of the behavior of politicians.

We expected that reevaluation of social norms would have a significant effect on judgments of both wrongness and punishment. Participants given the opportunity to cheat were predicted to be less harsh in their judgments.

Previous research (Keizer et al., 2008) has found that the environment in which an ethical decision is made can have a serious impact. Based on this finding, we would expect that mere exposure to a situation where unethical behavior goes unpunished decreases punitive judgments—possibly because the change in later judgments is based on participants' changing opinion of social norms.

METHOD

Participants

Sixty-two participants (28 male, 34 female) from the Harvard Decision Science Lab's human subject pool participated in this study for a maximum payment of \$15. The data from two participants were excluded as they did not complete all measures, leaving us with 60 participants (27 male, 33 female). The participants (not including the removed participants) ranged in ages from 18 to 40, and the average age was 26 years ($SD = 5.73$).

Subjects were guaranteed a payment of \$5 for participation, in addition to the money they could earn during the game—up to \$10 more.

Design and Procedure

Participants arrived at the lab and completed a consent form. They were then escorted to the main laboratory room and seated at individual cubicles.

Matrix number-matching task. When all participants were seated at their cubicles, they were given the following four test materials: a brown envelope containing \$10 (eight \$1 bills and four half-dollar coins); an empty white envelope; a worksheet with 20 matrices, each containing 12 numbers consisting of an integer and two decimals (e.g., 8.47); and a collection slip on which participants reported their performance and answered questions about their gender and age.

After all subjects were given the test materials and instructions were fully explained, a 5-min timer was started and subjects were told to begin. Participants had the full 5 min to find two numbers per matrix that added up to 10. They earned 50 cents for every matrix correctly solved.

Groups of participants were previously randomly assigned to either the confirmation or no confirmation condition. Condition determined the conclusion of the number-matching task, as follows:

1. Confirmation: After the 5 min had passed, participants were told to stop the number-matching task. They were asked to transfer their unearned money to the white envelope. The experimenter checked how many matrices each participant had solved correctly, and confirmed that score on the collection slip. The experimenter then confirmed that the correct amount had been left in the white envelope, and collected this envelope as well.
2. No Confirmation: At the end of the 5 min, participants were told to stop the number-matching task. They were asked to complete the collection slip, including reporting how many problems they solved and how much money they earned. Participants were then instructed to move their unearned money to the white envelope, to throw away their work-sheets in a recycling bin in the back of the room, and to give their collection sheets and unearned money to the experimenter on their way out. All instructions were given at once so that subjects could easily discern that their scores would not be checked. The experimenter made no effort to ensure that they reported their scores accurately.

Previous studies have found that subjects tend to solve an average of seven matrices when their performance is checked and report an average of 12 when given the opportunity to self-report (Gino et al., 2009). Often, this task is used to examine levels of dishonesty in response to different conditions, but this requires us to clandestinely review participants' answer sheets without their knowledge. In this study we did not examine how much individuals stole, nor did we compare average reported scores, due to a lab policy of no deception.

Ethical scenario judgments. Finally, participants were presented with a measure of punitive judgment. The questionnaire presented them with 10 scenarios of public officials who have done something that some people might consider to be corrupt or unethical. The scenarios come from previous research by Welch and Peters (1977). The scenarios were designed to provide a variety of examples of questionable behavior, providing both multiple types of corruption and multiple levels of severity. For example, a public official using his influence to get a friend or relative admitted to law school is relatively acceptable, whereas a legislator accepting a large campaign contribution in return for voting "the right way" on a legislative bill is almost always considered corrupt.

Participants were asked to judge on a 7-point Likert scale (1–7) both how wrong they considered each scenario to be, and how much punishment the official deserved, measures based on previous research in punitiveness (Cushman, 2008). Scenarios presented a number of types of unethical behavior (e.g., personal financial gain, using influence to gain favors for family or constituents, etc.), and behavior varied from mildly to severely unethical. Wrongness and punishment judgments were combined for each participant, leading to a sum score for each type of judgment (one aggregate wrongness score, and one aggregate punishment score). Only these aggregate scores were considered.

Following the study, participants were given a debriefing form and the flat fee (\$5). Participation took approximately 30 min.

RESULTS

A comparison of confirmation/no-confirmation condition and punitiveness/wrongness scores yielded significant results. Subjects in the no-confirm condition, that is, those who were able to cheat, rated the politicians' questionable actions as both less wrong ($M_{\text{confirm}} = 53.42$, $SD_{\text{confirm}} = 7.97$; $M_{\text{no-confirm}} = 47.89$, $SD_{\text{no-confirm}} = 9.46$), $t(58) = -2.43$, $p = .018$, and rated them as deserving less punishment ($M_{\text{confirm}} = 46.83$, $SD_{\text{confirm}} = 10.70$; $M_{\text{no-confirm}} = 40.86$, $SD_{\text{no-confirm}} = 8.78$), $t(58) = -2.36$, $p = .02$ (Figure 1). Cohen's effect size value suggested moderate practical significance for both wrongness ($d = 0.639$) and punitiveness ($d = 0.623$) comparisons.

Scores of punitiveness and wrongness were highly correlated, Pearson's $r(60) = .77$, $p < .001$ (see Table 1). A significant effect of sex was also found. Female participants rated the actions of the politicians as more wrong ($M_{\text{male}} = 47.33$, $SD_{\text{male}} = 9.20$; $M_{\text{female}} = 52.36$, $SD_{\text{female}} = 8.29$), $t(58) = 2.23$, $p = .03$, but there was no significant difference between how much men and women thought they should actually be punished ($M_{\text{male}} = 41.85$, $SD_{\text{male}} = 9.84$; $M_{\text{female}} = 44.39$, $SD_{\text{female}} = 10.05$), $t(58) = .98$, $p = .3$.

There was a significant correlation between the age of participants and wrongness scores, with younger participants rating the actions in the vignettes as more wrong, but no significant correlation between age and punitiveness (see Table 1).

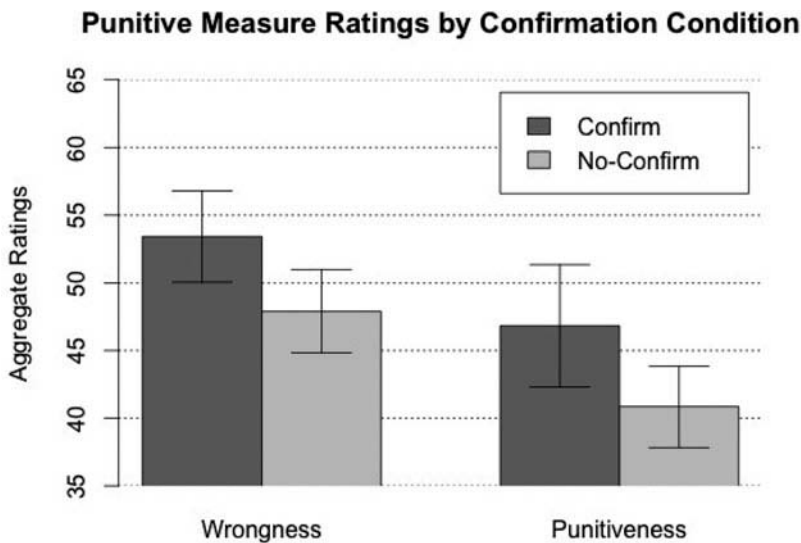


FIGURE 1 Aggregate scores of judgments of punitiveness and wrongness in the two confirmation conditions.

Note. Error bars represent 95% confidence interval for the mean.

TABLE 1
Zero-Order Correlations (Pearson's *r*)

	<i>Age</i>	<i>Wrongness Scores</i>	<i>Punitive Scores</i>
Age	1	—	—
Wrongness scores	−0.31**	1	—
Punitive scores	−0.15	0.77***	1

p* < .01. *p* < .001.

DISCUSSION

Both age and sex were found to have a significant effect on judgments of wrongness, with young participants and female participants rating the vignettes as more wrong. No similar effect was found, however, for judgments of punitiveness. Although it is not our aim to evaluate this relationship, a comparison to Cushman’s (2008) model suggests that, under his theory, younger and female participants may consider belief of the actor as more salient when making these judgments.

Confirmation and Ethical Judgments

The main finding of this study is that confirmation condition affects judgments of punishment and of wrongness. To put it in specific terms, being put in a situation where cheating is possible caused participants to later rate questionable behaviors less harshly.

It is not entirely clear what part of the no-confirmation condition is responsible for the change in judgments: actual cheating behavior, the suspicion that others are cheating, or the observation on the part of the participant that cheating would not be stopped. It may be, for example, that although actual cheaters’ judgments were significantly changed, those who were given the opportunity to cheat, but did not cheat, had judgments in line with the confirmation group—in this situation, behavior would have caused the later modification in judgment. Or it may be that the only factor needed to decrease punitiveness was the mere opportunity of acting unethically, whether or not participants actually took that opportunity. The two cases would be indistinguishable based on the data we have. A study in which actual levels of cheating were measured, however, might be able to shed light on this question.

We have considered three possible mechanisms that might be driving this effect—social norms, empathy, and cognitive dissonance.¹ All are plausible based on past research, and further studies should be conducted to determine which are responsible for the effect observed.

The first possible mechanism is that of social norms. It has been found that when participants are reminded of moral standards (Mazar et al., 2008a), they appear not to cheat at all. The reminder may serve to adjust their “baseline of morality” upward, so that when they are later

¹ An alternative explanation suggested by a reviewer is that the participants may have minimized the importance of their unethical actions. However, the literature on minimization is limited, and there is currently no reason to suspect that minimizing, if present, would be extended to judgments of others.

faced with the choice to cheat, they cannot cheat at all without updating their self-concept. The same thing may be occurring in this experiment—when cheating is permissible, their baseline of morality may drop, causing them to rate the vignettes less critically. Similarly, we have seen that participants will cheat based on what they believe is normal for their social group (Gino et al., 2009). In the current study, their perception of what is a normal ethical standard may have been revised downward, just as in Gino et al. (2009), which made them less harsh when later rating the behavior of others. Because they were just in a situation in which unethical behavior was not punished, or even frowned upon, they might then see such behavior as more acceptable or normal, a perception that shows itself during their later judgments.

Alternately, the change in judgment may be caused by empathy felt by the participant toward the politicians in the vignettes, who may be identified as fellow cheaters. Individuals can come to identify with each other very quickly on the basis of seemingly minor factors (Stürmer et al., 2006), and it may be that, having just cheated themselves, they identify with the elected officials presented during the judgment task. Even participants who did not cheat may still identify in some way, for they most likely considered cheating and can say to themselves, if not in so many words, “I was faced with a hard decision, and so was this person. I understand what they went through, even if they made the wrong choice.” This too may lead to the participant identifying with the person being judged. The same study found that participants were more helpful to even recently acquired and spurious ingroups. In our case, participants may be rating these individuals less harshly as a result of wanting to help them, due to their empathic response resulting from perceiving them as members of an ingroup. Certainly if the politicians were real, a less punitive judgment might be helpful to them. It has also been found (Gault & Sabini, 2000) that empathy causes participants to prefer victim-support solutions over punitive solutions, so it may be that empathy reduces punitiveness in general.

Finally, it may be that cognitive dissonance (Festinger, 1957) was responsible for the change in judgment. Cognitive dissonance, the feeling of discomfort caused by holding conflicting views, causes individuals to change their beliefs so that less dissonance is experienced (Aronson & Carlsmith, 1963; Festinger & Carlsmith, 1959). For participants who cheated on the number-matching task, there would be a contradiction between their desire to be harsh in their judgments and their knowledge of recently having made unethical decisions. As a result, they may have changed their views toward unethical behavior to make their attitude align more closely with their recent behavior. Although we did not intend to examine cognitive dissonance, we admit that it is a possible mechanism for the observed effect.

Fortunately, simple experiments could be conducted to examine which of these effects might have caused the observed change in judgments. It is of course possible that any combination of the three might be responsible, and there may be other mechanisms we have not considered.

Policy Implications

The findings of this study could play an important part in explaining how elected officials who are involved in scandals manage to win reelection (e.g., Richard Daley, Marion Barry, William J. Jefferson). If the social norms interpretation of these findings is correct, and the changes in judgment are caused by a reevaluation of “normal” standards for ethical behavior, then areas with prevalent unethical behavior would be more likely to elect or reelect an official regardless of

previous unethical behavior. Working to reduce crime in an area—more specifically, perceived crime—may actually lead to the election of more ethical officials, and to the ousting of corrupt officials. It may be that an effective method of fighting corruption at a governmental level is to fight crime—or even signs of crime—at a street level.

Under the social norms interpretation of these results, an effective method of combating unethical behavior in any situation is simply changing the perceived level of acceptable ethical behavior; this would be accomplished by either downplaying the level of unethical behavior or broadly promoting examples of good behavior. Under this interpretation, perception of the ethical behavior of others is the factor that individuals consider when making ethical decisions. Changing that perception would be a relatively easy method by which one could impact the ethical behavior of any group. Given this social norms interpretation, our findings seem to provide support for broken windows theory (Kelling & Wilson, 1982), which suggests that signs of disorder such as graffiti and litter induce further disorderly, and potentially unethical, behavior (Keizer et al., 2008).

It is possible that our results were not caused by a reevaluation of baseline of morality—our social norms interpretation. As already discussed, empathy as “fellow cheaters” for the elected officials being judged is a possible explanation as well. If empathy were the cause of our results, the main implications would be that empathy and identification with an individual causes judgments of that individual’s behavior to be less harsh. This is in line with previous findings (Gault & Sabini, 2000; Stürmer et al., 2006). It is important to remember, of course, that it is possible that both empathy and perception of social norms played a role in creating the effect we observed.

Although our measure of punitive judgment uses elected officials, we expect that similar results would be found for judgments of individuals who are not elected officials. We have no reason to think that people judge elected officials using different systems than they would use to judge nonofficials, so although exact judgments may differ, the effect should remain the same.

Difficulties in Reported Earnings

Although we did not use reported earnings in our analyses, reports were collected as part of the experiment. These reports were in many ways surprising. Although we expected that self-reports of earnings would be accurate, this was clearly not the case. Some participants reported impossible earnings (\$50) or reported earning amounts that did not match their self-reported score (reported one matrix solved, \$10 earned).

Also, examination of returned money in the white envelopes showed that there was a difference in reported earnings and actual earnings. We could not tie return envelopes to specific participants, but comparison of averages showed that more was being taken than was reported. Although we had expected to find a difference between performance and reported earnings in the no-confirm condition, a separate score for actual earnings came as a surprise. As an example, we had expected that a participant who completed 10 matrices (and thus should have earned \$5) might take \$7, and report the same. What we found was that subjects would complete 10 matrices (\$5), report \$7 earned, and take \$9 from the envelope (again, numbers are an example and not indicative of an actual instance). This means that reported scores were not an accurate measure of cheating.

Previous research using this task (e.g., Gino et al., 2009; Mazar et al., 2008a; Zhong et al., 2010) has reported no such discrepancies. It is important to note that previous research has

primarily used college students, whereas this study drew from the Harvard Decision Science Lab's human subject pool, which is much more diverse in both age and socioeconomic background. The finding that this group appears to cheat in a seriously different manner than previously studied student populations suggests that research on all-student populations may not fully generalize to ethical behavior among other groups. The fact that at least seven of our 60 participants (11.7%) cheated to the greatest extent possible (took all \$10), compared to only five of 791 (0.6%) in Mazar et al. (2008a) suggests that these types of experiments need to be repeated with more diverse populations, as behavior may be found to differ significantly.

The discrepancies found in this study do not refute previous findings, though they do suggest that the findings may not generalize as well as was suspected. The behavior of our more general population compared to student populations can be interpreted in a number of ways already suggested in the literature. It is possible that, for example, the general population may have a broader range of dishonesty within which they do not have to update their self-concept, or that more of them have cheated to such a degree that the "what-the-hell" effect is now dominating their behavior (Mazar, Amir, & Ariely, 2008b).

ACKNOWLEDGMENTS

A file containing the raw study data and the R code used to perform analysis may be found at http://thedata.harvard.edu/dvn/dv/Ethical_Situations/. We thank the Harvard Decision Science Laboratory for the use of facilities and support of this research. We thank Mark Edington, Sarah Hirschfeld-Sussman, SoYon Rim, and Joanna Morris for their help. Thanks also go out to Francesca Gino for providing us with the matrix task used in the experiment.

REFERENCES

- Aronson, E., & Carlsmith, J. M. (1963). Effect of the severity of threat on the devaluation of forbidden behavior. *Journal of Abnormal and Social Psychology*, 66, 584–588.
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2, 412–414.
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76, 169–217.
- Cushman, F. A. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108, 353–380.
- Festinger, L. (1957). *A theory of cognitive dissonance* (Vol. 2). Stanford, CA: Stanford University Press.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203–210.
- Gault, B. A., & Sabini, J. (2000). The roles of empathy, anger, and gender in predicting attitudes toward punitive, reparative, and preventative public policies. *Cognition and Emotion*, 14, 495–520.
- Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science*, 20, 393–398.
- Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, 322, 1681–1685.
- Kelling, G. L., & Wilson, J. Q. (1982, March). Broken windows. *The Atlantic*, p. 29.
- Mazar, N., Amir, O., & Ariely, D. (2008a). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45, 633–634.
- Mazar, N., Amir, O., & Ariely, D. (2008b). More ways to cheat—expanding the scope of dishonesty. *Journal of Marketing Research*, 45, 651–653.

- Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin*, 32, 943–956.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 96–102.
- Welch, S., & Peters, J. (1977). Attitudes of U. S. state legislators toward political corruption: Some preliminary findings. *Legislative Studies Quarterly*, 2, 445–463.
- Zhong, C. B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science*, 21, 311–314.