

# Personality and Social Psychology Bulletin

<http://psp.sagepub.com/>

---

## **Self-Interest Bias in Moral Judgments of Others' Actions**

Konrad Bocian and Bogdan Wojciszke

*Pers Soc Psychol Bull* 2014 40: 898 originally published online 17 April 2014

DOI: 10.1177/0146167214529800

The online version of this article can be found at:

<http://psp.sagepub.com/content/40/7/898>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[Society for Personality and Social Psychology](http://www.spsociety.org)

**Additional services and information for *Personality and Social Psychology Bulletin* can be found at:**

**Email Alerts:** <http://psp.sagepub.com/cgi/alerts>

**Subscriptions:** <http://psp.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Jun 24, 2014

[OnlineFirst Version of Record](#) - Apr 17, 2014

[What is This?](#)

# Self-Interest Bias in Moral Judgments of Others' Actions

Personality and Social Psychology Bulletin  
2014, Vol. 40(7) 898–909  
© 2014 by the Society for Personality and Social Psychology, Inc  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0146167214529800  
pspb.sagepub.com



Konrad Bocian<sup>1</sup> and Bogdan Wojciszke<sup>1</sup>

## Abstract

The automatic and affective nature of moral judgments leads to the expectation that these judgments are biased by an observer's own interests. Although the idea of self-interest bias is old, it has never been directly tested with respect to the moral judgments of other individuals' behaviors. The participants of three experiments observed other individuals' counternormative behavior (breaking a rule or cheating for gain), which was judged as immoral. However, this judgment became much more lenient when the observers gained from the observed behavior. All three studies showed that the influence of self-interest on moral judgments was completely mediated by the observer's increased liking for the perpetrator of the immoral acts but not by changes in mood. When the participants were induced to dislike the perpetrator (in a moderation-of-process design), the self-interest bias disappeared. Implications for the intuitionist approach to moral judgment are discussed.

## Keywords

self-interest, moral judgment, egotistic bias, ethics

Received November 21, 2013; revision accepted March 10, 2014

Morality is simply the attitude we adopt towards people whom we personally dislike.

—Oscar Wilde

The psychology of moral judgments has experienced a major change in the last two decades following the discovery that moral judgments do not have to rely on conscious, language-based reasoning but are frequently based on affect-laden intuitions that emerge quickly and automatically (Haidt, 2001; Haidt & Kesebir, 2010). This result has opened up prospects for new research and inspired the current study. We hypothesize a self-interest bias in moral judgments, such that other individuals' counternormative behaviors are perceived as less immoral when they serve the perceiver's own interests, whereas identical behaviors are considered more immoral when they are irrelevant to their own interests. We present three experiments demonstrating this bias and discern between its two explanations—mood congruence and attitude formation.

and those instant perceptions are difficult to inhibit or change by conscious efforts. Such judgments are generated without intention and occur even when the judges cannot explain them (Haidt, 2001) or are unable to do so because they are preverbal infants and toddlers (Hamlin, 2013). Changes in emotional states influence moral judgments even when factual information regarding a judged act is held constant. An experience of unrelated disgust increases the harshness of moral condemnation (Schnall, Haidt, Clore, & Jordan, 2008), and an experience of unrelated joy counteracts the normal flash of negative affect caused by the classical footbridge dilemma and leads to the neglect of deontological considerations when solving this dilemma (Valdesolo & DeSteno, 2006). Understandably, such affective influences are stronger in information-uncertain conditions, in which individuals lack relevant information for a particular situation (Van den Bos, 2003). Impairments of brain areas related to emotions—such as the ventro-medial prefrontal cortex and amygdala—impair moral judgments (Damasio, 1994), and healthy individuals with a low activation of these areas tend to ignore deontological

## Self-Interest Bias in Moral Judgments

There is a variety of data showing that moral judgments not only involve thoughtful deliberations but also automatic intuitions heavily based on affect. People make moral judgments instantly (within 200–250 ms—Van Berkum, Holleman, Nieuwland, Otten, & Murre, 2009),

<sup>1</sup>University of Social Sciences and Humanities, Sopot Campus, Poland

### Corresponding Author:

Konrad Bocian, Department of Psychology, University of Social Sciences and Humanities, Sopot Campus, Polna 16/20, 81-745 Sopot, Poland.  
Email: kbocian1@swps.edu.pl

considerations when solving moral dilemmas (Greene, Nystrom, Engell, Darley, & Cohen, 2004).

Given this automatic and affect-laden nature of moral judgments, individuals are bound to be highly susceptible to egocentric biases. Individuals automatically develop egocentric perceptions and evaluations (intuitively presuming that “what is good or bad for me is generally good or bad”), and because moral judgments of what is fair or unfair are inherently affective, judgments are generally based on those egocentric evaluations (Epley & Caruso, 2004). These egocentric biases are automatic because their intensity depends on heuristic cues (such as an approving nodding of the perceiver’s head when making a judgment—Epley & Gilovich, 2001), and they increase under time pressure but decrease in the presence of incentives to make accurate judgments (Epley, Keysar, Van Boven, & Gilovich, 2004). Continuing this line of reasoning on egocentric ethic, we argue that individuals are prone to *the self-interest bias in the moral judgments of others*. Namely, individuals perceive another individual’s counternormative behavior as moral when it serves their own (the perceivers’) interests, although they perceive the identical behavior as immoral when it is irrelevant to those interests. This self-interest bias arises for the following reasons: (a) moral evaluations are automatically based on current affective states and (b) affective responses to stimuli are driven by their bearing on the perceivers’ self-interest, such that the stimuli promoting self-interest generate positive affect and stimuli dwarfing self-interest generate negative affect.

There is substantial support for the self-interest bias in judgments of one’s own behavior and in self–other comparisons. When individuals decide on fair wage allocations, they claim more money for themselves than other individuals for identical work (Messick & Sentis, 1983). Similar differences are observed in the perceptions of a fair settlement in hypothetical disputes, in which perceivers play the role of plaintiff or defendant (Loewenstein, Issacharof, Camerer, & Babcock, 1993). Many studies on moral hypocrisy (Batson & Thompson, 2001) have shown that individuals behave in a self-interested manner when moral behavior would incur even mild costs. Nevertheless, individuals pretend to be moral and occasionally believe this pretense (Batson, Thompson, & Chen, 2002). For example, in the classical study of Batson, Kobrynowicz, Dinnerstein, Kampf, and Wilson (1997), the participants had to assign experimental tasks to themselves and an unknown other participant. One of the tasks was difficult, uncomfortable, and devoid of financial benefits, whereas the remaining task was easy and offered a prospective financial benefit. Unsurprisingly, 16 of 20 participants assigned the positive task to themselves, and even when they were allowed to flip a coin to make a fair decision, 10 of 20 individuals flipped the coin; nevertheless, 9 individuals assigned the positive task to themselves (Batson et al., 1997). These results suggest that individuals struggle between the desire for selfish gain and reputational concern

to not be recognized by others as selfish, thus engaging in maneuvers called “moral masquerade” by Batson (2008) to explain how the participants of his experiments attempted to conceal their real and deeply hidden selfish intentions.

A more recent study examined the effect of two positive emotions, pride and gratitude, on moral hypocrisy in an Asian context. The results showed that among Asian samples, moral hypocrisy was approximately as high as in the previously mentioned American samples. However, pride increased moral hypocrisy, but gratitude reduced it (Tong & Yang, 2011). Another study demonstrated that whereas angry individuals easily turn into moral hypocrites, the experience of guilt eliminates moral hypocrisy, and the experience of envy can reverse it because individuals judge others less severely than they judge themselves (Polman & Ruttan, 2012). This upturned moral hypocrisy effect was described as moral hypercrisis in a study of power, in which moral hypocrisy was enhanced by high power roles and reversed by low power roles (Lammers, Stapel, & Galinsky, 2010).

Despite these empirical developments, there is virtually no direct, undisputable support for the self-interest bias in judgments of others’ behaviors, although some studies may be interpreted as supportive. Studies on moral hypocrisy observed that the selfishness of in-groups was evaluated as morally right, whereas the selfishness of out-groups was morally wrong (Valdesolo & DeSteno, 2007), although this discrepancy was eliminated in cognitive load conditions (Valdesolo & DeSteno, 2008). In another study, two groups of participants evaluated a politician whose program was either beneficial or harmful to retirees. Whereas the participating students did not observe any difference between the two politicians, the participating retirees considered the politician acting in their group’s interest to be moral but the one acting against these interests was immoral (Cislak & Wojciszke, 2006). However, self-interest was confounded with the in-group interests in all of these studies. Therefore, it is not obvious how much the results support the hypothesis that self-interest biases the moral judgments of others’ behaviors. Perhaps the closest support is the result that the organizational leaders did not disapprove of the unethical behavior of their group members when they benefited from this behavior themselves (Hoogervorst, De Cremer, & van Dijke, 2010), but this study only measured global disapproval, which could be based on a variety of reasons.

To summarize, although the idea of self-interest in moral judgments is old, direct support for this idea concerns perceptions of the self and one’s own behavior rather than others and their behavior. Much of the empirical research has shown that perceptions of the self and others differ in many respects (Abele & Wojciszke, 2014; Pronin, 2009), including distinct differences in moral perceptions (Wojciszke, 2005). Thus, studies on moral self-judgments cannot be observed as informative about the perceptions of other individuals’ behaviors. Therefore, the hypothesis that self-interest biases the moral judgments of others’ actual

behaviors remains untested despite its familiarity. Even less is known of the exact mechanism of this bias; we could not locate an empirical study that addressed the underlying mechanisms of the self-interest bias in the moral judgments of others' behaviors.

### *Possible Mechanisms of the Self-Interest Bias*

The self-interest bias when judging the morality of others may be explained in at least two manners: as an effect of mood congruency or attitude formation. We investigated each mechanism because one of the goals of the present study was to ascertain their empirical validity as mechanisms underlying the bias.

*The mood-congruency hypothesis* assumes that observers who gain from an actor's rewarding behavior experience a rise in positive affect, which leads to more positive perceptions of the actor. Such mood-congruency effects on person perception were originally shown by Forgas and Bower (1987) who observed that happy participants formed more favorable impressions of others and make more positive judgments and developed more positive memories than sad participants. Similar mood effects were observed for a wide variety of judgments, from relatively inconsequential judgments, such as estimating general satisfaction with life during a survey (Schwarz & Clore, 1983) or assessing whether a facial expression was genuine (Forgas & East, 2008), to highly consequential judgments, such as medical school admissions (Redlmeier & Baxter, 2009) or stock market investments (which are higher in good weather conditions—Hirshleifer & Shumway, 2003). Mood-congruency effects in judgments have been explained in two major manners. Spreading-activation models assume that the current mood activates information that is affectively congruent, and these activated concepts or memories are integrated into judgments (Forgas & Bower, 1987). Affect-as-information models assume that mood can be used as a separate piece of information that is integrated into judgments when it is considered relevant and is not rejected as invalid (Schwarz & Clore, 2007).

When these two accounts are compared, affect-as-information typically appears more valid. For example, perceivers' judgments about other individuals tend to be more positive in good than bad mood, but when perceivers' attention is driven to their experimentally induced mood, they are able to correct this mood-congruency effect when judging others (McFarland, White, & Newth, 2003). Similarly, the effect of feelings increases in identical conditions that favor the influence of highly accessible information, such as low processing capacity (Greifeneder & Bless, 2007) or low motivation to process information (Rothman & Schwarz, 1998). Nevertheless, there are also models that assume that both accounts of mood congruency are valid, although in different conditions (Forgas, 1995).

*The attitude formation hypothesis* assumes that observers who gain from an actor's rewarding behavior develop a positive attitude toward the actor because the latter is instrumental in goal attainment. Objects that are instrumental for goal attainment are automatically evaluated in a more positive manner than identical objects that are irrelevant for an active goal (Ferguson & Bargh, 2004). Positive affect induced by goal achievement may be transferred to the means associated with goal attainment, and this affective transfer is proportional to the associative strength (Fishbach, Shah, & Kruglanski, 2004). Notably, individuals who facilitate goal attainment are also better liked by the perceiver who evaluates them more positively, draws closer to them, and approaches them more readily (Fitzsimons & Shah, 2008). The latter results only concerned significant others, but we assume that attitudes toward strangers function in a similar manner because those individuals who are instrumental in goal attainment are more appreciated, independent of the level of their acquaintance or closeness to the perceiver (Converse & Fishbach, 2012). Such increases in appreciation can be particularly expected in judgments of moral character.

Moral judgments of an individual heavily influence the liking of that individual, more so than judgments of other equally positive qualities (e.g., competence). This influence is completely mediated by the perceived benevolence of this person (Wojciszke, Abele, & Baryla, 2009). This result is because the morality of other individuals is directly rewarding for the perceiver (whereas other qualities, such as competence, may only be rewarding on the condition of good intentions—Abele & Wojciszke, 2014). Numerous studies have shown morality and liking to be strongly associated because information on the morality of an individual increases the liking of that individual and liking an individual increases the expectations of that individual as moral (Wojciszke, 2005). Bidirectional associations are a hallmark of the automatic/impulsive system of information processing (opposed to the conscious/reflective system, which is based on propositions; Strack & Deutsch, 2004). Because we assume that self-interest bias results from automatic processes, this is an additional argument for the attitude formation hypothesis. Therefore, it can be predicted that individuals who act for the perceivers' interests are perceived as more moral (even if breaking norms) because the perceivers like them better.

### *Overview of the Present Studies*

Study 1 was a field experiment, in which the self-interest bias was probed in a real-life setting, and the participants were not aware of their participation in the psychological study. The study occurred in a university library, in which a librarian broke the rules and waived (or not) a fine for overdue books. Three hypotheses were tested:

**Hypothesis 1:** The librarian who acted for the participants' interests would be observed as more moral compared with the librarian who did not waive the fine.

**Hypothesis 2:** The former librarian would be more liked.

**Hypothesis 3:** Differences in moral judgments would be mediated by differences in liking.

Studies 2 and 3 were laboratory experiments, in which the participants were paired with confederates who cheated for gain. The participants observed and then rated this behavior, whereas their own interests were/were not involved. In addition, mood was measured before and after the cheating to distinguish between mood congruency and attitude formation as explanations of the self-interest bias. Finally, in Study 3, we manipulated the participants' attitudes toward the confederates to obtain stronger evidence for the attitude formation process.

## Study 1

Three groups of students participated in this field study, which occurred in their library. The first group consisted of students who returned overdue books, and the librarian broke the rules and waived their fine. The second group consisted of students with overdue books but the fine was not waived. The third (control) group consisted of students who returned their books on time. After leaving the library, all students were approached by a confederate that was conducting a seemingly unrelated survey that concerned the quality of services provided by the department, including the library and dean's office. The questions included ratings of efficiency, morality, and liking of the librarian and an employee of the dean's office.

We tested the hypotheses that the librarian acting for the participants' interests would be observed as more moral and likeable and that differences in moral judgments would be mediated by differences in liking. We also tested whether these differences in perceptions of the librarian would carry over to judgments of the dean's office employee. The mood-congruency hypothesis suggests such carry-over effects due to the diffused nature of (presumed) increases in mood resulting from self-interest fulfillment. However, the attitude formation hypothesis predicts no such effects because it assumes that the affective change mediating the self-interest bias in moral judgments is a person-bound liking for the perpetrator of the act in question.

## Method

**Participants.** Sixty undergraduate students (57 women;  $M$  age = 26.45 years) who visited a library of their university participated in this field experiment.

**Procedure and materials.** The study began when a student, who had to pay a fine for overdue books, went to the library. The librarian informed the participants regarding the amount

of the fine and then randomly either waived or imposed the fine. In this manner, one group of 20 students gained by keeping their money ( $M_{\text{gain}} = 15$  Polish zlotys;  $SD = 12.53$ ; approximately US\$4.53), whereas another group of 20 students lost money by paying the fine ( $M_{\text{loss}} = 12.1$  zlotys;  $SD = 8.42$ ; approximately US\$3.65; the difference between the two means was not significant,  $p = .396$ ). A control group consisted of another 20 students who visited the library and returned their books on time. Waiving the fine was the librarian's counternormative behavior and illicitly benefitted the student, whereas imposing the fine was a normative behavior that was harmful for the student.

When the students left the library, they were followed by a female confederate for 3 min and then approached and asked to participate in a short survey (see Appendix) concerning two university departments: the library and dean's office (in a random order). The confederate introduced herself as an employee of the marketing department to make the request more plausible and reduce suspicions regarding the real aim of the study. Because none of the participants expressed any suspicion regarding the true aim of the study or showed distress, we did not inform the participants that they had participated in a psychological study. Our decision was driven by a public opinion poll on law observance in Poland reported by Public Opinion Research Center. According to the report, 51% of Poles think that the majority of ordinary people in Poland often disobey the law, whereas only 44% of Poles think that usually people in Poland comply with the law (CBOS, 2013). This suggests that participants did not experience anything that was beyond their everyday experience.

The participants evaluated the infrastructure of the departments (e.g., chairs, lights, and the Internet speed) and quality of the personnel. Specifically, on 7-point scale (1 = *definitely no*; 7 = *definitely yes*), the participants rated how much they agreed that the employees were *efficient*, *fast*, *sensible*, *accurate*, *capable*, and *lazy* in their work, which were filler adjectives for the three items that measured moral evaluations: *honest*, *fair*, and *proper*. The latter adjectives were borrowed from Abele and Wojciszke (2007, Study 1) who had many trait names rated for morality-relatedness by a group of 20 raters. The average ratings for these three items were 4.65, 4.20, and 4.40 (on a scale ranging from -5 to 0 to 5), which showed that all three traits were regarded as highly moral. On a 7-point scale (1 = *strongly negative*, 7 = *strongly positive*), the students also rated their overall attitude toward the library and the dean's office staff (always the last contacted person). Because liking is a major constituent of interpersonal attitudes (as shown by Wojciszke et al., 2009), we will use the terms *overall attitude* and *liking* interchangeably.

## Results

**Moral evaluation and liking.** The three critical ratings of the librarian's and dean's office employees' actions were averaged and yielded the moral evaluation measures (Cronbach's

**Table 1.** Ratings of Morality and Liking of The Librarian's and The Dean's Office Employee as Function of Three Conditions of Study 1.

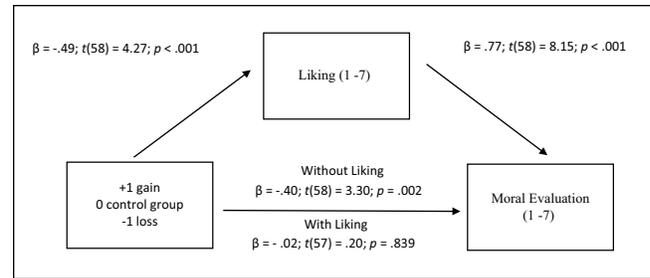
Condition	Moral evaluation	Liking
Target person: Librarian		
Gain	6.22 (0.80)	6.25 (0.79)
Control	5.92 (0.81)	5.80 (0.95)
Loss	5.30 (1.02)	4.70 (1.56)
Target person: Dean's office employee		
Gain	5.83 (.70)	5.95 (0.89)
Control	5.68 (0.88)	5.60 (0.99)
Loss	5.48 (1.02)	5.55 (0.99)

Note. Standard deviations are given in parentheses. Higher numbers mean more lenient moral evaluation and positive attitude (liking).

$\alpha = .94$  and  $\alpha = .89$ ). A one-factor ANOVA of moral evaluations of the librarian appeared significant,  $F(2, 57) = 5.57$ ;  $p = .006$ ;  $\eta^2 = .16$ , therefore, we moved to selected pair comparisons between the means. The loss condition yielded lower moral evaluations ( $M = 5.30$ ,  $SD = 1.02$ ) than the control condition ( $M = 5.92$ ,  $SD = 0.81$ ),  $t(38) = 2.11$ ,  $p = .041$ ,  $d = 0.67$ , 95% confidence interval (CI) = [0.02, 1.30] as predicted. Similarly, the loss condition yielded lower evaluations than the gain condition ( $M = 6.22$ ,  $SD = 0.80$ ),  $t(38) = 3.16$ ,  $p = .003$ ,  $d = 1$ , 95% CI = [0.33, 1.64]. The difference between the gain and control conditions was not statistically significant,  $p = .249$ . Moreover, these differences were only limited to judgments of the librarian because in the case of the dean's office employee, the overall one-factor ANOVA was not significant,  $p = .452$ . This result renders the explanation of the self-interest bias in terms of mood elevation improbable because the diffused nature of mood effects implies that they should influence all evaluative judgments, including those of individuals other than the perpetrator of the acts that served the observer's interest.

Similar analyses were performed for the liking judgments. Because the overall ANOVA on liking (attitude) was significant,  $F(2, 57) = 9.64$ ,  $p = .001$ ,  $\eta^2 = .25$ , pair comparisons were conducted for the overall attitude (liking) toward the librarian. Table 1 shows that the librarian was liked the least in the loss condition ( $M = 4.70$ ,  $SD = 1.56$ ) but to an upper degree in the control ( $M = 5.80$ ,  $SD = 0.95$ ),  $t(38) = 31.42$ ,  $p = .011$ ,  $d = 0.85$ , 95% CI = [0.19, 1.48] and gain conditions ( $M = 6.25$ ,  $SD = 0.79$ ,  $p = .001$ ),  $t(38) = 28.08$ ,  $p < .001$ ,  $d = 1.25$ , 95% CI = [0.55, 1.90]. The difference between the gain and control conditions was not statistically significant ( $p = .111$ ). For the dean's office employee, the difference between the means was again insignificant ( $p = .364$ ).

**Mediating the effect of liking.** To test our prediction that liking mediated the relation between self-interest and moral evaluation, we employed the mediational procedure (R. M. Baron & Kenny, 1986). Figure 1 shows that the direct effect of self-interest ( $\beta = .40$ ) was reduced to insignificant when liking



**Figure 1.** Liking as a mediator of the self-interest influence on moral judgment (Study 1).

was included in the analyses ( $\beta = .02$ ), and liking was a significant predictor of moral evaluation ( $\beta = .77$ ). We tested the overall significance of the indirect effect by devising a 95% CI, as advocated by Preacher and Hayes (2004). Because zero fell outside of the interval [0.2262, 0.6796], the indirect effect of self-interest on moral evaluation was significant (cf. MacKinnon, Fairchild, & Fritz, 2007).

## Discussion

Study 1 demonstrates the self-interest bias in moral judgments and the underlying mechanism. The counternormative behavior of another individual serving the observer's interest was judged as more moral than normative behavior, which produced deserved harm to the observer. The counternormative behavior also increased the liking of the actor, and this increased liking completely mediated the difference in moral judgments. No support was observed for the mood-congruence explanation of the self-interest bias: Inconsistent with this explanation, increases in judgments of the act perpetrator (the librarian) did not carry over to the judgments of another individual (a dean's office employee).

## Study 2

In this experiment, the participants observed and evaluated another individual (the actor) who cheated to win an eye-catching prize. In one condition, the observers' interests were involved (they obtained a similar prize if the paired actor won), in the second condition, the observers' interests were not involved (only the cheating actor could win the prize). To explore the mood-congruency hypothesis further, we also measured the participants' mood before and after they observed cheating.

## Method

**Participants.** Forty undergraduate students (35 women;  $M$  age = 26.87 years) participated in this study in exchange for course credit.

**Procedure and materials.** The participants entered the laboratory individually and were paired with another student

(a female confederate) who arrived 2 min later. Both students were informed that they would participate in a lab study on “perspective-taking”: One individual would be the actor and the other individual would be the observer (for exact instructions and materials see Appendix). All (actual) participants were assigned the observer position by a rigged drawing. The actors and observers were seated in separate cubicles and randomly assigned to a condition. In the self-interest involvement condition, we informed the participants that because of funding from a research grant, there were two prizes to win, one for each member of the team, whereas in the condition that did not involve self-interest, the participants were informed that only the actor could win the prize. The prize was a brand new iPod shuffle (one or two). The chance of winning the prize depended on the number of mazes solved by the actor, and each completed maze contributed one ticket for the lottery. The actor’s task was to solve as many online mazes as she could in 5 min, whereas the observer’s task was to monitor how the actor solved the mazes. The participants were informed that the mazes could be solved by clicking either “Auto-Solve” or “Path-Verify” but were explicitly prohibited from using these functions because it would be considered as cheating. This procedure was shaped after Gneezy, Niederle, and Rustichini (2003).

We measured mood using the Positive and Negative Affect Schedule scale (PANAS; Watson, Clark, & Tellegen, 1988) preadapted to Polish. We linked the actors’ and observers’ computers, thus allowing the observer to track how the actor was performing the task. To standardize the conditions, the participants observed a prerecorded movie, during which only the two first mazes were solved fairly and the other five were solved using the forbidden functions. This movie was started during a moment when the participants’ attention was diverted to another detail of the procedure. The participants watched the cheating actor (confederate) for 5 min and then answered questions identical to Study 1 (moral evaluation Cronbach’s  $\alpha = .93$ ). The participants also answered questions regarding satisfaction with their role and attention to the actors’ performance (no between-condition differences were observed). Finally, the participants’ mood was measured again with PANAS scale.

## Results

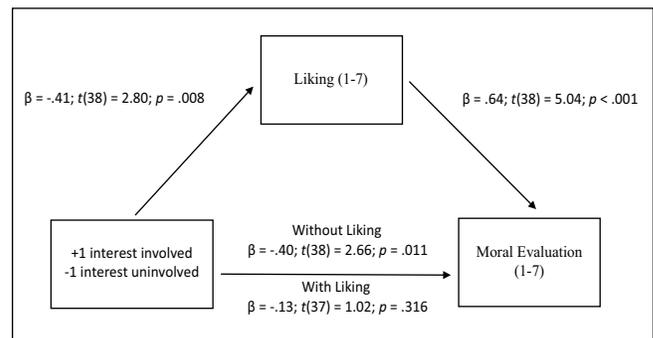
The participants who had an opportunity to win a reward together with the perpetrator were satisfied with their role as an observer to an identical extent as the participants without this possibility ( $M = 5.15$ ,  $SD = 1.10$  vs.  $M = 4.90$ ,  $SD = 1.02$ ,  $p = .461$ ). The number of mazes solved by the perpetrator did not differ between conditions ( $p = .509$ ) and there were no differences in feelings of anonymity ( $p = .683$ ) or attention devoted to observations of the perpetrator’s actions ( $p = .329$ ).

**Moral evaluations and liking.** The actors’ unethical behavior was evaluated as more moral when the participants’ interest

**Table 2.** Ratings of the Actors’ Morality, Liking, and Mood as Function of the (Observers’) Self-Interest Involvement Condition of Study 2.

Variable	Interest involved	Interest uninvolved
Moral evaluations	4.57 (2.25)	2.85 (1.81)
Liking	5.30 (1.21)	4.30 (1.03)
Initial mood	15.05 (6.30)	13.95 (6.30)
Final mood	10.05 (8.52)	10.60 (6.90)

Note. Standard deviations are given in parentheses. Initial and final mood indicates participants’ overall mood at the beginning and at the end of experiment. Higher numbers mean more lenient moral evaluation, more positive attitude (liking) and more positive mood (initial and final).



**Figure 2.** Liking as a mediator of the self-interest influence on moral judgments (Study 2).

was involved ( $M = 4.57$ ,  $SD = 2.25$ ) than when it was not involved ( $M = 2.85$ ,  $SD = 1.81$ ),  $t(38) = 2.66$ ,  $p = .011$ ,  $d = .84$ , 95% CI = [0.18, 1.47], as was predicted. A similar difference was also observed for the overall attitude toward the actor,  $t(38) = 2.80$ ,  $p = .008$ ,  $d = .89$ , 95% CI = [0.22, 1.52], which was higher in the condition that involved the observers’ interest ( $M = 5.30$ ,  $SD = 1.21$ ) than did not involve this interest ( $M = 4.30$ ,  $SD = 1.03$ ; see Table 2).

**Liking as a mediator.** The mediation analysis yielded similar results to those of Study 1. The direct effect of self-interest on moral judgment ( $\beta = .40$ ,  $p = .011$ ) was reduced to insignificant when the liking of the actor was included ( $\beta = .13$ ,  $p = .316$ ), and liking was a significant predictor of moral evaluation ( $\beta = .64$ ,  $p < .001$ ), as was observed in Figure 2. The indirect effect was significant because zero fell outside the appropriate 0.95 interval [0.1534, 1.1738].

**Mood.** The overall ANOVA on mood in a 2 (self-interest: involved vs. uninvolved)  $\times$  2 (repeated measures: before vs. after) design showed a main effect of the time of mood measurement,  $F(1, 38) = 10.64$ ,  $p = .002$ ,  $\eta^2 = 0.22$ . Independent of the condition, the participants’ overall mood decreased at the end of experiment ( $M_{\text{before}} = 14.50$ ,  $SD = 6.21$  vs.  $M_{\text{after}} = 10.32$ ,  $SD = 7.65$ ), thus suggesting that witnessing others’ cheating decreases the observer’s mood. This result was consistent with other data that showed that immoral acts of

others lead to negative emotional states in observers (Wojciszke & Szymkow, 2003). Neither the main effect of manipulation nor interaction between the manipulation and mood measurement time were significant ( $p = .881$  and  $p = .523$ , respectively; see descriptive statistics in Table 2). This pattern of mood change was not consistent with what would be expected if mood were the mediator of the influence of self-interest on moral judgments.

### Discussion

Study 2 replicated the results from Study 1 under controlled lab conditions with a different manipulation of self-interest, thereby increasing the external generalizability validity of our results. We again showed a mediating role of liking and observed no such role for mood. However, mood decreased as a result of witnessing immoral behavior.

### Study 3

In this experiment, we aimed to replicate the previous results, thus changing the manipulation of self-interest and provided stronger—experimental, not merely mediational—evidence for liking as the underlying mechanism of the self-interest bias. Adopting a “moderation-of-process” design (Spencer, Zanna, & Fong, 2005), we manipulated liking to show that self-interest bias only emerged when the immorally acting offender was likeable. Evaluations from other individuals are the key determinant of their attractiveness because the attraction is reciprocated, particularly in terms of affective responses (Montoya & Insko, 2008). Negative evaluations received from other individuals can inhibit attraction resulting from similarity (Aronson & Worchel, 1966), although negative evaluations are a strong antecedent of attraction (Montoya, Horton, & Kirchner, 2008). Therefore, we assumed that negative evaluations from the actor would disallow her liking, as has been observed in many studies (e.g., Hewitt, 1972; Insko et al., 1973; Sigall & Aronson, 1969), which, in turn, would inhibit increases in moral judgments of the actions serving the observer’s interests. Thus, we expected that this inhibition of liking would prevent the emergence of self-interest bias.

### Method

**Participants.** In all, 120 undergraduates (95 women;  $M$  age = 25.87 years) participated in this study in exchange for course credit. The participants were also offered an opportunity to earn additional money.

**Procedure and materials.** Study 3 was a replication of Study 2 with three exceptions. First, we involved the participants’ interests more directly and provided them an opportunity to collect up to US\$10 instead of having the opportunity to win a prize.

Second, the confederate was a different woman from Study 2, and the actor’s task consisted of solving 10 simple mathematic equations (e.g.,  $53 - 12 + 10 - 15 + 6 - 19 = ?$ ) in 5 min without external help. After five fairly solved equations, the confederate removed a cell phone from her handbag, turned on a calculator, and solved the remaining equations. The situation was observed by the participants sitting in a neighboring cubicle via webcam directed at the actor’s desk, which allowed them to see that the actor cheated (i.e., used a calculator). Simultaneously, the participants checked the actor’s answers with an answer key and paid US\$1 for each correct answer. Only the actor was paid in the uninvolved self-interest condition, whereas both the actor and observer were paid that amount each in the involved self-interest condition.

Third, we manipulated the attitudes toward the perpetrator (liking vs. disliking vs. no manipulation) and adopted seven personality traits from Willis and Todorov (2006; that is, attractive, trustworthy, aggressive, calm, decent, ambitious, and open for new experiences). Before the main task, the participants were informed that we wanted to explore person perception processes among newly acquainted individuals. Under this pretext, both the actors and observers responded to a questionnaire and rated one another’s seven traits on 7-point scales (1 = *definitely not*, 7 = *definitely yes*). The participants were then provided questionnaires that were allegedly completed by their partners and learned that the partner either liked (the ratings varied but all were above the midpoint) or disliked them (the ratings varied but all were below the midpoint). To examine how the extent to which the participants agreed with the actors ratings, we asked them to examine the ratings and rate on a 7-point scale (1 = *definitely not accurate*, 7 = *definitely accurate*) to what extent the ratings were accurate estimates of their personality. In the control condition, the participants were moved immediately to the main task without the liking manipulation. The remaining portions of the experiment were identical to Study 2, including the two measurements of mood by PANAS scale. At the end of the experiment, all participants were debriefed and thanked for their participation (for research scenario and materials see Appendix).

### Results

Manipulations of interest and liking did not affect the participants’ satisfaction with their role of an observer ( $p = .320$ ) or the number of solved equations ( $p = .889$ ), feelings of anonymity ( $p = .754$ ), and attention devoted to the observation of the perpetrator’s actions ( $p = .889$ ). The participants in the liking condition evaluated the perpetrator’s ratings regarding their personal traits as more accurate ( $M = 5.55$ ,  $SD = 0.71$ ) than participants in the disliking condition ( $M = 1.80$ ,  $SD = 1.00$ ),  $F(1, 76) = 383.75$ ;  $p < .001$ ;  $\eta^2 = .83$ , regardless of the interest manipulation, as was predicted. The latter result suggests that the liking manipulation was successful, as far as

**Table 3.** Ratings of the Actors' Morality and Mood as Function of the Self-Interest Involvement and Liking Condition of Study 3.

Condition	Interest involved	Interest uninvolved
<b>Morality</b>		
Liking	5.80 (1.57)	3.15 (1.68)
Disliking	3.57 (1.65)	2.98 (1.55)
Control	4.67 (1.74)	3.53 (2.40)
<b>Initial mood</b>		
Liking	12.30 (6.74)	14.30 (12.28)
Disliking	13.60 (6.09)	11.15 (8.75)
Control	17.70 (8.04)	13.40 (6.18)
<b>Final mood</b>		
Liking	14.75 (11.52)	12.15 (13.33)
Disliking	8.40 (9.09)	5.40 (7.27)
Control	17.70 (9.48)	9.35 (8.30)

Note. Standard deviations are given in parentheses. Initial and final mood indicates participants' overall mood at the beginning and at the end of experiment. Higher numbers mean more lenient moral evaluation and more positive mood (initial and final).

the accuracy judgments can be observed as a proxy measure of liking.

**Moral evaluations.** A 2 (interest involved: yes vs. no)  $\times$  3 (liking: liking vs. disliking vs. control) ANOVA on moral evaluations (Cronbach's  $\alpha = .93$ ) showed a significant main effect of interest. Moral judgments in the self-interest involved condition were higher ( $M = 4.68$ ,  $SD = 1.87$ ) than in the condition not involving self-interest ( $M = 3.22$ ,  $SD = 1.90$ ),  $F(1, 114) = 19.86$ ,  $p < .001$ ,  $\eta^2 = .15$ , thus replicating the results of Studies 1 and 2. The results also showed a significant main effect of the liking manipulation,  $F(2, 114) = 4.71$ ,  $p = .011$ ,  $\eta^2 = .08$ , although post hoc analyses showed that participants in the disliking condition were harsher in their moral evaluations ( $M = 3.27$ ,  $SD = 1.61$ ) than those in the liking ( $M = 4.47$ ,  $SD = 2.09$ ),  $p = .003$ , or control conditions ( $M = 4.10$ ,  $SD = 2.15$ ),  $p = .041$ , and the difference between the last two conditions was not significant.

Most importantly, the predicted interaction between interest and liking appeared significant,  $F(2, 114) = 3.58$ ,  $p = .031$ ,  $\eta^2 = .06$ . In the liking condition, participants whose interest was involved showed higher moral evaluations ( $M = 5.80$ ,  $SD = 1.57$ ) than those whose interest was not involved ( $M = 3.15$ ,  $SD = 1.68$ ),  $t(38) = 5.15$ ,  $p < .001$ ,  $d = 1.63$ , 95% CI = [0.89, 2.31]. However, in the disliking condition, moral judgments were similar for participants whose interest was both involved ( $M = 3.57$ ,  $SD = 1.65$ ) and uninvolved ( $M = 2.98$ ,  $SD = 1.55$ ),  $t(38) = 1.15$ ,  $p = .256$ , as shown in Table 3.

This interaction confirms that liking accounts for leniency in moral judgments when self-interest is involved and that self-interest bias is eliminated when the perpetrator of immoral acts (the actor) is disliked. In the control condition, the difference between the interest conditions appeared significant  $t(38) = 1.71$ ,  $p = .045$  (one-tailed test), thus replicating the results of our former studies.

**Mood.** A 2 (interests)  $\times$  3 (liking)  $\times$  2 (moment of mood measurement) ANOVA revealed a main effect of mood,  $F(1, 114) = 9.03$ ,  $p = .003$ ,  $\eta^2 = .07$ , such that the participants' overall mood decreased at the end of the experiment ( $M_{\text{before}} = 13.74$ ,  $SD = 8.37$  vs.  $M_{\text{after}} = 11.29$ ,  $SD = 10.65$ ). This effect was not constrained by any manipulation, and no other significant effects were revealed (see the descriptive statistics in Table 3). This pattern of mood change was not consistent with what would be expected if mood were a mediator.

## General Discussion

The idea of self-interest bias is certainly not novel, but it has curiously never been studied systematically with regard to moral judgments of other individuals' behaviors. Many researchers have theorized that because of its automatic and uncontrollable nature, self-interest bias is powerful enough to overcome morally relevant concerns such as norm-based deliberations (Epley & Caruso, 2004) or professional obligations (Moore & Loewenstein, 2004). Nevertheless, this bias appears to be taken for granted rather than studied. The absence of empirical development is most likely responsible for the fact that the notion of self-interest bias virtually disappeared from recent monographs on moral judgments (Haidt, 2012; Mikulincer & Shaver, 2012) and the chapter on morality in the latest edition of *Handbook of Social Psychology* (Haidt & Kesebir, 2010).

The present research is the first line of studies that systematically show self-interest bias in the moral judgments of other individuals' counternormative behaviors. As shown in Studies 2 and 3, such behavior is judged as immoral, but when the observers gain from the behavior, their judgments become much more lenient, up to changing the valence of their perceptions to positive. Study 1 observed a similar—although more lenient—pattern for a milder transgression of breaking a library rule (which, judging by the general mean, was not observed as immoral).

The moral evaluations in Study 1 were higher than those in the next two studies and can be explained in two manners. On one hand, the norm violation in Study 1 (disregarding library regulations) can be perceived as less immoral than the norm violations in Studies 2 (cheating for a prize) and 3 (cheating for cash). To test this hypothesis, we described the three behaviors to a sample of students ( $N = 57$ ) from an identical population as the participants of the main studies. We asked the participants how *honest*, *fair*, and *proper* each behavior was (on 7-point scales ranging from  $-3 = \text{definitely not}$  to  $3 = \text{definitely yes}$ ). The norm violation from Study 1 was judged as less immoral ( $M = -1.84$ ,  $SD = 1.04$ ) than the norm violations from Study 2 ( $M = -2.47$ ,  $SD = 1.15$ ),  $t(56) = 4.31$ ,  $p < .001$ , or Study 3 ( $M = -2.50$ ,  $SD = 1.10$ ),  $t(56) = 4.47$ ,  $p < .001$ , as was predicted. For the norm violations of Studies 2 and 3, the difference between the means was insignificant ( $p = .715$ ). However, only the norm violation benefited the participant but not the actor in Study 1, and the norm violation appeared as an error of omission rather

than commission. Wrongful omissions of other individuals are judged less harshly than wrongful commissions (Anderson, 2003; J. Baron & Ritov, 2004; Cushman, Young, & Hauser, 2006; DeScioli, Christner, & Kurzban, 2011), and this result could have also contributed to the difference between the moral evaluations in Study 1 compared with Studies 2 and 3.

A variety of research on moral hypocrisy (Batson, Thompson, Seufferling, Whitney, & Strongman, 1999) or cheating (Mazar, Amir, & Ariely, 2008; Shu, Gino, & Bazerman, 2011) has shown that individuals retain a clear conscience and beliefs in their own morality after behaving unethically. The present studies showed that a similar change (objective) of dishonesty or rule breaking into (subjective) honesty also occurs to the behaviors of other individuals when the observer personally gains from this behavior.

The original contribution of the present study was that we studied two explanations of the self-interest bias, mood congruency and attitude formation, which shed new light on the mechanisms underlying the self-interest bias. According to the first explanation, the observers who gained from an actor's rewarding behavior experienced an increase in positive mood, which led to more positive perceptions of the actor (and other objects). No support was observed for this explanation. Study 1 cast some doubt on the role of mood because no increases in judgments of individuals other than the actor (the librarian) were observed, which should be expected because of the diffused nature of mood and the manner in which it influences judgments (Forgas, 1995). Study 2 ruled mood even more because it showed a significant decrease in mood after observing other individuals who cheated, and this result was replicated in Study 3, which also established a convincing case for the alternative explanation. The decreases in mood could be attributed to our participants becoming bored throughout the studies, but such an explanation was hardly plausible for the Study 3 participants who observed how the actor earned an increasing amount of money. Nevertheless, the PANASscale includes one item that explicitly measures excitement; analyses of this particular item did not show significant decreases in excitement in Study 2 or 3.

The attitude formation hypothesis fared better—all three studies observed that the observers who gained from an actor's behavior developed a positive attitude toward the actor, and these increases in liking completely mediated increases in moral judgments. Moreover, Study 3 observed causal evidence for the role of attitudes. When the observers were induced to dislike an actor, the self-interest bias disappeared, and the actor remained immoral in the eyes of those who disliked her, as was predicted by Oscar Wilde cited at the outset. This result has some disconcerting implications, which is typical for the self-interest bias. When self-interest is involved, individuals may reframe others' counternormative behaviors as virtuous. For example, a librarian waiving a fee may be reframed as an authority figure who exhibits prosociality toward a beleaguered student—a generous act of

human-to-human kindness. A student who cheats on a task during a psychology experiment may be viewed as someone displaying a small, virtuous act of rebellion against a powerful, capricious-seeming authority figure (the experimenter). It is possible that the participants were not viewing the actor's act as immoral. Moreover, in both cases, the participant and actor were conspiring together. Thus, a small conspiratorial in-group was created. Individuals obviously tend to perceive behaviors that favor the in-group as moral (Cislak & Wojciszke, 2006; see Leach, Bilali, & Pagliaro, 2014, for a review) and expect (objectively) immoral favors from in-groups (Foddy, Platow, & Yamagishi, 2009). Therefore, future studies should assess participants' thoughts and feelings regarding an actor and their relations with the actor to elucidate the exact mechanism of rationalization and reframing counternormative behavior as morally acceptable and "good." The present studies suggest that cherishing positive attitudes toward other individuals—because of cooperation, similarity, trust, and otherwise laudable qualities—increases the risk of being biased and forming overly positive impressions of their behavior as serving the common interest of the perceiver and behavior perpetrators. A possible method to prevent individuals from being biased by self-interest, for example, in business or external audit organizations, would be to bond their interests with individuals outside of their group because the perceived dissimilarity predicts disliking (Singh & Ho, 2000).

Together, these results shed new light on the intuitionist approach to moral judgments (Greene, 2007; Haidt, 2007). Morally relevant behavior of other individuals does not lead to a uniform affective response. We observed the opposite response: that identical behavior may lead to decreases in mood and increases in the liking of the actor, which suggests that an affective response may be nuanced and multifaceted. The gains acquired by the immorality of other individuals possibly do not translate into increases in declared mood because such declarations would be a blatant acknowledgment of an illicit joy. However, the response of liking another individual is most likely observed as anchored in that person, not the self, and therefore may be admitted freely. Such dislocation of illegitimate affects from a perceiver to an observer after immoral behavior opens interesting prospects for future research.

Other possible routes for research include factors mitigating the presently shown self-interest bias. There are good reasons to believe that this bias can be reduced by accountability because accountability induces more thoughtful information processing and diminishes nearly all biases (Lerner & Tetlock, 1999). Self-awareness may serve a similar function because it increases allegiance with social norms and personal values (Batson et al., 1999; Wicklund, 1975). Primes of social control, such as a picture of human eyes (Bateson, Nettle, & Roberts, 2006), may also decrease self-interest bias because social control increases norm-consistent action and judgment. Finally, the bias may wane over time. This result was suggested by the fact that perceivers'

appreciation for helpful individuals peak before help is complete (Converse & Fishbach, 2012) but then decrease over time (Flynn, 2003) after the tasks are complete and the relevant goals are deactivated.

## Conclusion

Using the observation of real behavior, we showed that self-interest biases the moral judgments of other individuals' unethical actions. Because the interests of different observers can diverge tremendously, this result contributes to a better understanding of why moral judgments are often controversial and disputable (despite objectivity claims of moral judges). The self-interest bias shows that moral judgment can be a matter of liking (or disliking) rather than merely following principles. In turn, this suggestion raises a serious concern regarding business ethics because scandals could be avoided if company leaders and employees stop turning a blind eye to the unethical practices of their well-liked colleagues (Gino, Moore, & Bazerman, 2009). Hopefully, a better understanding of how self-interest biases moral judgments could help us create environments that are less conducive to unethical actions.

## Acknowledgments

We want to thank Kasia Dehling, Barbara Makrenek, Aleksandra Pawlowicz, Agnieszka Rosner, and Beata Blazejewicz for their help in data gathering and Wieslaw Baryla for his invaluable statistical knowledge.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Foundation for Polish Science [Grant MASTER/MISTRZ].

## References

- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*, 751-763.
- Abele, A. E., & Wojciszke, B. (in press, 2014). Communal and agentic content in social cognition: A double perspective model. *Advances in Experimental Social Psychology*.
- Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin, 129*, 139-167.
- Aronson, E., & Worchel, P. (1966). Similarity vs. liking as determinants of interpersonal attractiveness. *Psychonomic Science, 5*, 157-158.
- Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes, 94*, 74-85.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters, 2*, 412-414.
- Batson, C. D. (2008). Moral masquerades: Experimental exploration of the nature of moral motivation. *Phenomenology and the Cognitive Sciences, 7*, 51-66.
- Batson, C. D., Kobryniewicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology, 72*, 1335-1348.
- Batson, C. D., & Thompson, E. R. (2001). Why don't moral people act morally? Motivational considerations. *Current Directions in Psychological Sciences, 10*, 54-57.
- Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology, 83*, 330-339.
- Batson, C. D., Thompson, E. R., Seufferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology, 77*, 525-537.
- CBOS. (2013, March 5). *Opinions about observance of law and work of justice system in Poland*. Retrieved from <http://cbos.pl/EN/publications/reports.php>
- Cislak, A., & Wojciszke, B. (2006). The role of self-interest and competence in attitudes toward politicians. *Polish Psychological Bulletin, 37*, 203-212.
- Converse, B. A., & Fishbach, A. (2012). Instrumentality boosts appreciation: Helpers are more appreciated while they are useful. *Psychological Science, 23*, 560-566.
- Cushman, F., Young, L., & Hauser, M. D. (2006). The role of reasoning and intuition in moral judgments: Testing three principles of harm. *Psychological Science, 17*, 1082-1089.
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York, NY: Putnam.
- DeScioli, P., Christner, J., & Kurzban, R. (2011). The omission strategy. *Psychological Science, 22*, 442-446.
- Epley, N., & Caruso, E. M. (2004). Egocentric ethics. *Social Justice Research, 17*, 171-187.
- Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Divergent processing of self-generated and experimenter-provided anchors. *Psychological Science, 12*, 391-396.
- Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*, 327-339.
- Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: The effects of goals pursuit on automatic evaluation. *Journal of Personality and Social Psychology, 87*, 557-572.
- Fishbach, A., Shah, J. Y., & Kruglanski, A. W. (2004). Emotional transfer in goal systems. *Journal Experimental Social Psychology, 40*, 723-738.
- Fitzsimons, G. M., & Shah, J. Y. (2008). How goal instrumentality shapes relationship evaluations. *Journal of Personality and Social Psychology, 95*, 319-337.
- Flynn, F. J. (2003). What have you done for me lately? Temporal adjustments to favor evaluations. *Organizational Behavior and Human Decision Processes, 91*, 38-50.

- Foddy, M., Platow, M. J., & Yamagishi, T. (2009). Group-based trust in strangers: The role of stereotypes and expectations. *Psychological Science, 20*, 419-422.
- Forgas, J. P. (1995). Mood and judgment: The Affect Infusion Model (AIM). *Psychological Bulletin, 116*, 39-66.
- Forgas, J. P., & Bower, G. H. (1987). Mood effects on person perception judgments. *Journal of Personality and Social Psychology, 53*, 53-60.
- Forgas, J. P., & East, R. (2008). How real is that smile? Mood effects on accepting or rejecting the veracity of emotional facial expressions. *Journal of Nonverbal Behavior, 32*, 157-170.
- Gino, F., Moore, D. A., & Bazerman, M. H. (2009). See no evil: When we overlook other people's unethical behavior. In R. M. Kramer, A. E. Tenbrunsel, & M. H. Bazerman (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments* (pp. 241-263). New York, NY: Psychology Press.
- Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics, 118*, 1049-1074.
- Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences, 11*, 322-323.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*, 389-400.
- Greifeneder, R., & Bless, H. (2007). Relying on accessible content versus accessibility experiences: The case of processing capacity. *Social Cognition, 25*, 853-881.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998-1002.
- Haidt, J. (2012). *The righteous mind. Why good people are divided by politics and religion*. New York, NY: Pantheon Books.
- Haidt, J., & Koseoff, S. (2010). Morality. In S. Fiske & D. Gilbert (Eds.), *Handbook of social psychology* (5th ed., pp. 797-832). Hoboken, NJ: Wiley.
- Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers: Evidence for an innate moral code. *Current Directions in Psychological Science, 22*, 183-193.
- Hewitt, J. U. (1972). Liking and the proportion of favorable evaluations. *Journal of Personality and Social Psychology, 22*, 231-235.
- Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance, 58*, 1009-1032.
- Hoogervorst, N., De Cremer, D., & van Dijke, M. H. (2010). Why leaders not always disapprove of unethical follower behavior: It depends on the leader's self-interest and accountability. *Journal of Business Ethics, 95*, 29-41.
- Insko, C. A., Thompson, V. D., Stroebe, W., Shaud, K. F., Pinner, B. E., & Layton, B. D. (1973). Implied evaluation and the similarity-attraction effect. *Journal of Personality and Social Psychology, 25*, 297-308.
- Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy: Moralizing in reasoning, immorality in behavior. *Psychological Science, 21*, 737-744.
- Leach, C. W., Bilali, R., & Pagliaro, S. (in press, 2014). Groups and morality. In J. Simpson & J. F. Dovidio (Eds.), *APA handbook of personality and social psychology, vol. 2: Interpersonal relationships and group processes*. Washington, DC: American Psychological Association.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*, 255-275.
- Loewenstein, G., Issacharof, S., Camerer, C., & Babcock, L. (1993). Self-serving assessments of fairness and pretrial bargaining. *Journal of Legal Studies, 22*, 135-159.
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593-614.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*, 633-644.
- McFarland, C., White, K., & Newth, S. (2003). Mood acknowledgment and correction for the mood-congruency bias in social judgment. *Journal of Experimental Social Psychology, 39*, 483-491.
- Messick, D. M., & Sentis, K. (1983). Fairness, preference, and fairness biases. In D. M. Messick & K. S. Cook (Eds.), *Equity theory: Psychological and sociological perspectives* (pp. 61-94). New York: Praeger.
- Mikulincer, M., & Shaver, P. R. (Eds.). (2012). *The social psychology of morality*. Washington, DC: American Psychological Association.
- Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships, 25*, 889-922.
- Montoya, R. M., & Insko, C. A. (2008). Toward a more complete understanding of the reciprocity of liking effect. *European Journal of Social Psychology, 38*, 477-498.
- Moore, D. A., & Loewenstein, G. (2004). Self-interest automaticity, and the psychology of conflict of interest. *Social Justice Research, 17*, 189-202.
- Polman, E., & Ruttan, R. L. (2012). Effects of anger, guilt, and envy on moral hypocrisy. *Personality and Social Psychology Bulletin, 38*, 129-139.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717-731.
- Pronin, E. (2009). The introspection illusion. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 41, pp. 1-67). Burlington, MA: Academic Press.
- Redlmeier, D. A., & Baxter, S. D. (2009). Rainy weather and medical school admission interviews. *Canadian Medical Association Journal, 181*, 933.
- Rothman, A. J., & Schwarz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin, 24*, 1053-1064.
- Schnall, S., Haidt, J., Clore, G., & Jordan, A. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*, 1096-1109.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513-523.
- Schwarz, N., & Clore, G. L. (2007). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 385-407). New York, NY: Guilford.

- Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin, 37*, 330-349.
- Sigall, H., & Aronson, E. (1969). Liking for an evaluator as a function of her physical attractiveness and nature of the evaluations. *Journal of Experimental Social Psychology, 5*, 93-100.
- Singh, R., & Ho, S. Y. (2000). Attitudes and attraction: A new test of the attraction, repulsion, and similarity-dissimilarity asymmetry hypotheses. *British Journal of Social Psychology, 39*, 197-211.
- 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.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220-247.
- Tong, E. M. W., & Yang, Z. (2011). Moral hypocrisy: Of proud and grateful people. *Social Psychological & Personality Science, 2*, 159-165.
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science, 17*, 476-477.
- Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science, 18*, 689-690.
- Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44*, 1334-1338.
- Van Berkum, J. J. A., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science, 20*, 1092-1099.
- Van den Bos, K. (2003). On the subjective quality of social justice: The role of affect as information in the psychology of justice judgments. *Journal of Personality and Social Psychology, 85*, 482-498.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.
- Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 233-275). New York, NY: Academic Press.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science, 17*, 592-598.
- Wojciszke, B. (2005). Morality and competence in person- and self-perception. *European Review of Social Psychology, 16*, 155-188.
- Wojciszke, B., Abele, A. E., & Baryla, W. (2009). Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology, 39*, 973-990.
- Wojciszke, B., & Szymkow, A. E. (2003). Emotions related to others' competence and morality. *Polish Psychological Bulletin, 34*, 135-142.