

Other Papers

Polish Psychological Bulletin
 2016, vol. 47(3) 366–372
 DOI - 10.1515/ppb-2016-0043

Konrad Bocian*
 Wiesław Baryła*
 Bogdan Wojciszke*

When Dishonesty Leads to Trust: Moral Judgments Biased by Self-interest are Truly Believed

Abstract: Research has shown that cheating is perceived as immoral when it serves the cheater's interests, though it can be seen as moral when it serves the interests of the perceiver. However, are such biased moral judgments real, or are they merely lip service? To answer the question of whether biased moral judgments actually inform behavior, the authors asked participants to observe a confederate who either cheated for money or did not cheat, which benefited either the confederate alone or both the confederate and the participating observer. Then, participants evaluated the confederate and, finally, played a one shot trust game with her. Cheating influenced moral judgments and decreased behavioral trust, but this only occurred when self-interest was not involved. When self-interest was involved, participants showed no significant differences in trust levels, independent of whether the confederate had cheated or not. Implications for the dual process theory in moral psychology are discussed.

Key words: moral judgments, self-interest bias, cheating, trust

“Self-interest makes some people blind, and others sharp-sighted”.

Francois de La Rochefoucauld

There is increasing agreement that moral judgments are a joint product of automatic and controlled processes. Moral intuitions are fast and continuous; therefore, they play a greater role than controlled deliberations, which are slower and require motivation and generous conditions in order to appear (Conway & Gawronski, 2013; Green, 2007; Haidt, 2007). Moral intuitions are affective in nature, and because self-interest can be a strong source of affect, this allows moral judgments to be biased by self-interest. Such biases have been frequently shown in regard to moral judgments of one's own behavior as *moral hypocrisy* – the tendency to evaluate moral transgressions as less objectionable when they are enacted by the self, rather than other persons (Batson, Thompson, Seuferling, Whitney, & Strongman, 1999). A similar self-interest bias has been recently reported in regard to interpersonal perceptions. Bocian and Wojciszke (2014a) showed that others' counter-normative behavior (breaking a rule or cheating

for money) leads the perceiver to make more positive moral evaluations when he or she materially benefits from this behavior, when compared to a situation where only the perpetrator (the rule-breaker) benefits. The same studies showed that this change in judgment is entirely mediated by affective responses toward the perpetrator. A perpetrator whose behavior benefits observers becomes well-liked by them, and this liking leads to heightened moral evaluations of his or her behavior (though changes in overall mood do not serve as a mediator). Research also showed that people are unaware of self-interest bias because they omit their emotional responses when forecasting their moral judgments (Bocian & Wojciszke, 2014b).

On the other hand, several studies have shown that self-interest can lead to deliberation process. For example, hypocritical judgments require cognitive resources (which suggests their deliberative nature, Valdesolo & DeSteno, 2008) and are easily extended to in-groups (Valdesolo

* SWPS University of Social Sciences and Humanities, Sopot, Poland

& DeSteno, 2007). Studies on public goods games showed that deliberation leads to more selfish behavior (Rand, Greene, Nowak, 2012) which corresponding with results from dictator game where participants in a high-load condition were more generous and often offered fair allocation that subject in low-load condition (Schulz, Fischbacher, Thöni Utikal, 2014). Undoubtedly, self-interest can inflate moral judgments of both one's own and others' behavior. However, an important question about such judgments remains unanswered: Do people truly believe their biased moral evaluations? This question is especially important in regard to interpersonal evaluations because these evaluations drive social behavior. Therefore, do people act upon their biased moral judgments and show trust toward highly evaluated persons, even if these moral evaluations have been inflated by the fact that these persons cheated (for the perceiver's interest)? Trusting such cheaters would be a paradoxical response because cheating typically leads to decreases in trust. Numerous anthropological observations of small-scale societies show that cheaters are avoided and excluded from social exchanges (Baumard, 2010), and experimental games show that cheaters are punished (even at a cost), which is mediated by negative emotions and followed by increased cooperation (Fehr & Gächter, 2002). Cheaters are also better remembered, and distrust in them is resistant to extinction when the information about their dishonesty is found to be untrue (Suzuki, Honma & Suga, 2013). All of this evidence suggests that in order to protect themselves, people should remain wary of cheaters and should not trust them.

On the other hand, recent research has shown that cheaters are remembered very well when their immoral behavior creates costs for the observers, but memory of cheaters is much poorer when their immorality creates benefits for the observers (Bell, Schain, & Echterhoff, 2014). Moreover, people believe in the objectivity of their own moral judgments and assume that these judgments reflect true facts. One's belief in the objectivity of his or her own moral judgments is nearly as high as one's belief in factual judgments (such as the fact that Mars is smaller than Venus) and much higher than one's belief in judgments about social conventions (such as remaining silent during a lecture) or preferences (believing that classic music is better than rock music; Goodwin & Darley, 2008). Due to one's belief in the objective nature of moral judgments, these judgments should inform behavior, and people who are considered to be moral should be trusted, even if the observer's moral judgments have been inflated by cheating that benefited the observer. Failing to act upon such judgments would lead to cognitive dissonance, which is unpleasant and preemptively avoided (Cooper, 2007). Not acting upon these judgments would also be seen as a hypocritical discrepancy, which acts as a threat to self-integrity and is avoided by engaging in behaviors that are consistent with one's own judgments, as shown by many studies (e.g., Stone, 2011).

Thus, there are strong arguments for either increases or decreases in trust to occur in persons who have previously

cheated and benefited the observers. In order to solve this question empirically, we performed a 2×2 experiment where participants first observed a confederate who either cheated for money or did not cheat, which benefited either the confederate alone or both the confederate and the participating observer. Then, observers judged the morality and their liking of and perceived closeness to the confederate and, finally, played a standard one shot trust game with her. For control reasons, mood was also measured both at the beginning and the end of the experiment.

Method

Participants and Design

Bocian and Wojciszke (2014a) reported medium ($r = .26$, $n = 40$, Study 3) and large ($r = .39$, $n = 40$, Study 2) relations between self-interest involvement and moral judgments, the average effect size was .33. Power analyses were conducted using G*Power (Faul, Erdfelder, Lang & Buchner, 2007). Results suggest that, given alpha of .05 and power of .80, a sample of 52 total participants would be required to detect an effect size of .33. A total of 120 participants were included in the present study, half of whom were randomly allotted to the cheating condition group. As such, the present sample size is sufficiently powered for replication of this medium effect of the self-interest involvement on moral judgments. On the other hand, sample of 60 participants in the no self-interest involvement condition is more than sufficient for the expected large effect of honesty on moral judgments. The present sample size is also sufficiently powered for medium to large effects on trust. According to the table 3 in Fritz and MacKinnon (2007) a total sample of 71 to 90 participants is needed to achieve .80 power on $p = .05$ when testing for mediation with medium sizes of α and β paths. Therefore the sample of 120 participants seems sufficient to testing if moral judgments mediated the influence of cheating on behavioral trust.

One hundred twenty undergraduates (92 women; mean age = 25.24 years) participated in the study in exchange for a \$3 show-up reward and course credit. The participants were also offered an opportunity to earn additional money. Seven participants were excluded: four did not believe the cover story, and the other three guessed the actual purpose of the study. This left a total of 113 participants, who were randomly assigned to a condition within the 2 (cheating: yes vs. no) $\times 2$ (self-interest: involved vs. not involved) design.

Procedure and measures

The procedure was an extension of the method that was developed by Bocian and Wojciszke (2014a, Study 3). Participants entered the laboratory individually and were paired with another student (a female confederate) who arrived two minutes later. After a brief introduction, both were informed that they would participate in a lab study on "perspective-taking". One person would be an actor (confederate), and the other would be an observer. Due to

the rigged drawing, all of the participants were assigned the observer position, and afterward, the confederate and the participant as observer were escorted to separate cubicles. The actor's task was to solve ten simple mathematic equations (e.g., $53 - 12 + 10 - 15 + 6 - 19 = ?$) in five minutes, without external help. Depending on the condition, after four of the equations were easily solved, the confederate either took a cellphone out of her handbag, turned on a calculator and solved the remaining equations (cheating condition) or simply solved all of the equations without the use of the calculator (no cheating condition). The situation was observed by the participant, who was sitting in a neighboring cubicle, via a webcam that was directed at the confederate's desk, which allowed participant to see that the confederate cheated (i.e., used a calculator) or did not cheat. We told participants that camera observation is a part of the "perspective taking" in the study which means that other student (confederate) knows she will be observed. We also told participants that their task is to observe carefully how other student is executing the mathematical task. Simultaneously, the participants checked the confederate's answers with an answer key, and paid \$0.50 either only the confederate (the uninvolved self-interest condition) or both the participant and the confederate (the involved self-interest condition) for each correct answer.

First, we measured the participants' mood by using the PANAS scale (Watson, Clark & Tellegen, 1988), which had been pre-adapted to Polish; then, after the rigged drawing, we asked participants about their satisfaction with their role. Afterward, we linked the actors' and observers' computers, which allowed the observer to track the actor's performance on the task. When five minutes had passed, the actors' and observers' computers were disconnected, and the participants rated on 7-point scale (1 = definitely no; 7 = definitely yes) how much they agreed that their partner completed the task *efficiently*, *quickly*, *sensibly*, *accurately*, *capably* and *lazily*, which were filler adjectives for the three items that measured moral evaluations: *honestly*, *fairly*, and *properly* (Cronbach's alpha = .91). The participants also rated their overall attitude toward the confederate on a 7-point scale (1 = strongly negative; 7 = strongly positive), and for control purposes, they answered questions about their attention to the actor's performance.

During the second stage of the study, we asked participants to play a one-shot trust game, in order to measure their trustworthiness toward partner (Berg, Dickhaut, & McCabe, 1995), in which they could pass to their partner (confederate) \$0 to \$3 from their show-up reward. The participants learned that the experimenter would triple the amount of money that was moved to the partner, who would decide how much money would be returned to the participant. Participants put money in an envelope and gave it back to the experimenter, who handed it over to their partner (i.e. confederate). Although the confederate pretended to deliberate over how much money to return to the participant, she always returned half of the tripled amount. We measured the participants' feelings of closeness toward the partner by using the Inclusion of Other

in the Self (IOS) Scale (Aron, Aron, & Smollan, 1992). By the time that the participants finished filling out the questionnaires, the experimenter had already handed them the envelope that contained the money that was returned from the partner. Finally, the participants' mood was measured again with the PANAS scale, and they were then extensively debriefed for any suspicions and thanked for their participation.

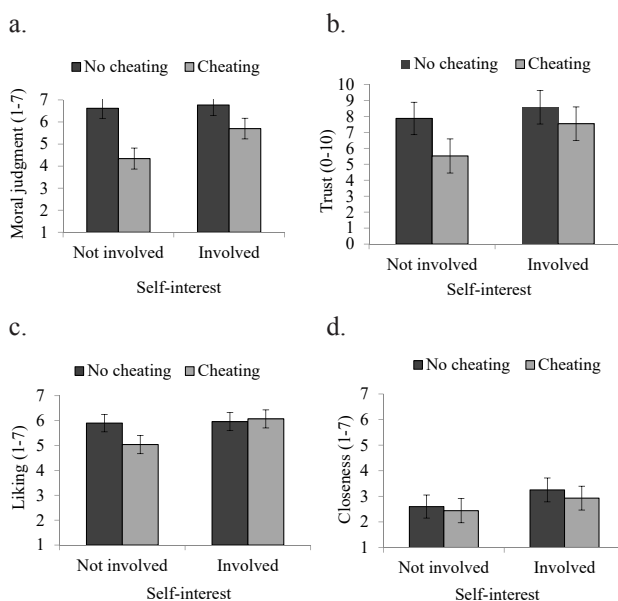
Results

Main analyzes

Moral judgments were subjected to an analysis of variance based on the 2 (self-interest: involved vs. not involved) \times 2 (cheating: no vs. yes) design. This analysis revealed a strong main effect of cheating $F(1,109) = 49.80$, $p < .001$, $\eta_p^2 = .314$. Moral judgments were lower in the cheating condition ($M = 5.03$, $SD = 1.87$) than in the non-cheating condition ($M = 6.69$, $SD = .42$). The main effect of self-interest involvement was also statistically significant, $F(1,109) = 9.88$, $p = .002$, $\eta_p^2 = .083$, although its interpretation was constrained by a significant interaction between the two factors, $F(1,109) = 6.28$, $p = .014$, $\eta_p^2 = .054$. As seen in Figure 1a, the interaction indicates that cheating decreased moral judgments both when self-interest was involved, $t(30.861) = 3.57$, $p = .001$, $d = .93$, 95% CI (.37, 1.52) and when self-interest was not involved, $t(28.232) = 5.93$, $p < .001$, $d = 1.57$, 95% CI (.90, 2.23); however, in the latter condition, the decrease was much larger. Therefore, cheating led to higher moral judgments when self-interest was involved ($M = 5.70$, $SD = 1.55$) than when self-interest was not involved ($M = 4.34$, $SD = 1.95$), $t(53) = 2.83$, $p = .007$, $d = .79$, 95% CI (.21, 1.31). These data show a clear self-interest bias in moral judgments of others' dishonest behavior, therefore replicating the results of Bocian and Wojciszke (2014a) and complementing the data on moral hypocrisy, which show that one's own transgressions are evaluated less harshly than the identical transgressions of others (Batson et al., 1999).

A similar two-factor analysis of variance was performed on the behavioral trust index (the amount of money that was invested in Polish zlotys, 0–10). This analysis revealed two main effects. Trust was higher in the involved self-interest condition ($M = 8.05$, $SD = 2.40$) than in the condition that did not involve self-interest ($M = 6.75$, $SD = 3.41$), $F(1,109) = 6.50$, $p = .012$, $\eta_p^2 = .056$. Trust appeared lower in the cheating condition ($M = 6.54$, $SD = 3.30$) than in the non-cheating condition ($M = 8.21$, $SD = 2.49$), $F(1,109) = 10.05$, $p = .002$, $\eta_p^2 = .084$ (cf., Figure 1b). However, a simple effect analysis showed that the decrease in trust, that was induced by cheating, was only reliable when self-interest was not involved ($M_{\text{no cheating}} = 7.87$, $SD = 2.73$ vs. $M_{\text{cheating}} = 5.52$, $SD = 3.72$), $t(55) = 2.73$, $p = .008$, $d = .73$, 95% CI (.19, 1.26). When self-interest was involved, the difference between the cheating condition ($M = 7.54$, $SD = 2.51$) and the no cheating condition ($M = 8.57$, $SD = 2.20$) failed to reach statistical significance, $t(54) = 1.64$, $p = .106$, $d = .44$, 95% CI (-.09, .97). Finally, in the cheating condition,

Figure 1. Moral judgments (a), behavioral trust (b), liking (c), and perceived closeness (d) as a function of cheating and the involvement of participants' self-interest



participants showed higher trust when they benefited from the cheating than when they did not benefit from the cheating, $t(45.44) = 2.37, p = .023, d = 0.63, 95\% \text{ CI } (.09, 1.17)$. Thus, cheating decreased trust, but this only occurred when self-interest was not involved. When self-interest was involved, participants showed the same high level of trust, independent of whether the confederate cheated or not. In other words, our participants truly believed in their moral judgments that were inflated by their self-interest and acted upon these judgments by trusting the cheater (confederate) with 80% of their money that was available to use.

The self-interest by cheating analysis of variance was repeated for the measures of liking of the confederate and perceived closeness to the confederate. As seen in Figure 1c, for liking, the dominant finding was an interaction between the two factors, $F(1,109) = 6.63, p = .010, \eta_p^2 = .060$. Cheating reduced liking for the perpetrator, but this only occurred when self-interest was not involved, $t(42,346) = 2.82, p = .007, d = .75, 95\% \text{ CI } (.21, 1.28)$. When self-interest was involved, the confederate was highly liked, independent of whether she cheated or not. As a result, a cheater who acted for the participant's interest was more liked than a cheater who only acted for her own interest ($M_{\text{interest involved}} = 6.07, SD = 0.81$ vs. $M_{\text{interest not involved}} = 5.04, SD = 1.37$), $t(41.958) = 3.38, p = .002, d = 0.91, 95\% \text{ CI } (.34, 1.47)$. For perceived closeness, the analysis only yielded one statistically significant effect: a main effect of self-interest involvement, $F(1,109) = 5.72, p = .019, \eta_p^2 = .050$. As seen in Figure 1d, a confederate who acted for the participant's interest was perceived as closer to the participant than one who acted for her own interests alone ($M_{\text{interest involved}} = 3.09, SD = 1.25$ vs. $M_{\text{interest not involved}} = 2.53, SD = 1.26$).

Finally, we subjected the mood measure to a 2 (self-interest involvement: involved vs. not involved) \times 2 (cheating: yes vs. no) \times 2 (time of measurement: end of study vs. beginning of study) ANOVA, with repeated measurements on the last factor. This analysis revealed two statistically significant effects: a main effect of the time of the measurement, $F(1,109) = 5.00, p = .027, \eta_p^2 = .044$, and an interaction between cheating and the time of measurement, which heavily constrained the interpretation of the main effect, $F(1,109) = 5.75, p = .018, \eta_p^2 = .050$. In the no cheating condition, mood was higher at the end ($M = 1.78, SD = 1.08$) than at the beginning ($M = 1.34, SD = 1.02$) of the study, $t(57) = 3.05, p = .003, d = .57, 95\% \text{ CI } (.29, 0.83)$. In the cheating condition, mood was the same at the end ($M = 1.57, SD = 1.08$) and at the beginning ($M = 1.58, SD = 0.83$) of the study, $t < 1$. Because there was no further interaction with self-interest involvement, this suggests that participation in the study was pleasant for the participants, but observing the confederate's dishonesty eliminated this feeling of pleasantness.

Moderated mediation

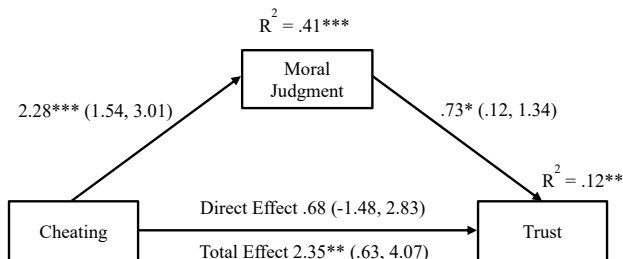
Logically, the influence of cheating on behavioral trust should be mediated by moral judgments. However, because cheating did not influence trust when self-interest was involved, we expected this mediation to be moderated by self-interest. In order to test this expectation, we conducted a moderated mediation analysis by using the PROCESS macro for SPSS, as suggested by Hayes (2013, Model 58). As seen in Table 1 and Figure 2, self-interest involvement moderated the effect of cheating on moral judgments,

Table 1. Outcome of the PROCESS macro (Model 58) that was used to test whether self-interest involvement moderates the mediation of moral judgments on the effect of cheating on trust

Variable	B	SE	P
Mediator (Moral judgments) Variable Model			
Constant	7.15	1.20	.000
Cheating (Independent Variable)	-.11	.75	.885
Self-interest involvement (Moderator)	-2.54	.76	.001
Cheating \times Self-interest involvement	1.19	.48	.014
Dependent (Trust) Variable Model			
Constant	10.02	3.96	.013
Moral judgments (Mediator)	-.37	.64	.559
Cheating (Independent Variable)	.79	.62	.208
Self-interest involvement (Moderator)	-4.21	2.23	.062
Moral judgments \times Self-interest involvement	.54	.36	.135

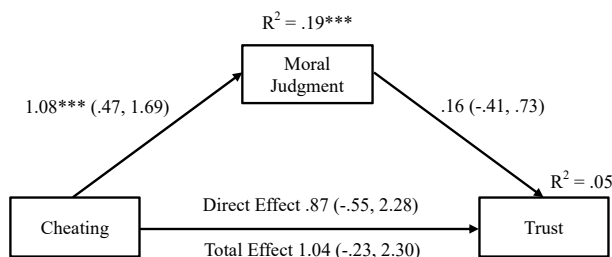
Figure 2. Mediation of the influence of cheating on behavioral trust by moral judgments when self-interest is (a) or is not involved (b)

a. Self-interest not involved



*** $p < .001$; ** $p < .01$; * $p < .05$

b. Self-interest involved



*** $p < .001$; ** $p < .01$; * $p < .05$

as well as the effect of moral judgments on behavioral trust. Providing evidence for moderated mediation, the index of moderated mediation was statistically significant ($b = 1.440$, $SE = .75$, Boot 95% CI: .10 to 2.95), and the effect of cheating on moral judgments was significantly higher in the condition that did not involve self-interest ($b = 2.277$, $SE = .34$, 95% CI: 1.61 to 2.94) than in the condition that involved self-interest ($b = 1.083$, $SE = .34$, 95% CI: .41 to 1.75). Additionally, the effect of moral judgments on trust was significantly higher (and significant) in the non-cheating condition ($b = 1.625$, $SE = .67$, Boot 95% CI: .30 to 2.93) than in the cheating condition ($b = .185$, $SE = .41$, Boot 95% CI: -.45 to 1.21), where it was not significant.

Clearly, moral judgments mediated the influence of cheating on behavioral trust, but this only occurred when self-interest was not involved. When self-interest was involved, this mediation disappeared, mainly because there was no difference in trust to be mediated: when participants benefited from the confederate's actions, they showed high trust in both the cheating and the non-cheating confederates. Finally, although the liking of the confederate and perceived closeness to the confederate were influenced by the conditions (Figures 1c and 1d) and correlated with trust, none of these variables (nor mood) served as a mediator of the influence of the conditions on trust.

In summary, other persons' dishonesty decreases both moral judgments and behavioral trust in these persons, but this only occurs when the perceiver's interests are not involved. When the perceiver's interests are involved, and

the perceiver benefits financially from others' dishonesty, the other person is judged more leniently, and there is no decrease in behavioral trust in the perpetrator, in the likeability of the perpetrator, or in the perceiver's closeness to the perpetrator.

Discussion and Limitations

Present findings unequivocally show that people trust in a cheater is not undermined due to self-interest bias and believe that their moral judgments are true, so they act upon these judgments. This meshes well with other results showing that people are unaware of the fact that self-interest influences their moral judgments (though they are aware of the influence of norms, Bocian & Wojciszke, 2014b). However, we are mindful that they are at least two alternative explanations for our results.

One can argue that participants in our study perceived the motives of the cheater differently in the situation when confederate benefited herself vs. benefited both herself and the participant. In the uninvolved cheating condition, the confederate could be perceived as acting selfishly, whereas in the involved condition participants could see her as a person who is motivated by altruistic motives. A recent study showed that prosocial lies increased the willingness to pass money in a trust game (Levine, Schweitzer, 2014) which could also explain presented findings, beyond intuitive account. Further studies could check how participants perceive confederate actions by asking them about her motives. We could assume that participants would see the confederate's behavior as cheating, but they will ascribe her different motives. However, this can be true only if participants will know that confederate knows she is observed. Therefore, this factor should be examined in the further studies.

The present method where participants know that the confederate knows she is observed could have created a shared identity. In other words, one can assume that the self-interest involved condition created a situation of interdependency between a participant and the confederate. In turn, participants could perceive the cheater as a person who did not cheat the trustee but cheated the third person and increased the payoff for both involved parties. A recent study found strong effects of collaboration on cheating (Weisel & Shalvi, 2015) showing that when people have to decide of whether to remain honest or join forces in collaborative cheating, they strongly opt for engaging in a corrupt collaboration. Further studies should rule out this alternative explanation by making participants fully independent from the confederate by assuring participants that confederate does not know that she is observed.

There is also a possibility that cheating vs. no cheating condition could indirectly imply the confederate's lack of competence and that is why she resorted to immoral behavior as the only way to complete the task. In order to clarify this potential confound, we averaged filler adjectives ratings (*efficient, quick, sensible, accurate, capable, lazy*) creating an agency evaluation measure

(Cronbach's $\alpha = .71$). We also conducted an overall ANOVA on competence and morality judgments in a 2 (confederate cheating: no vs. yes) \times 2 (judgment content: agency vs. morality) design with repeated measurement on the second factor. This analysis yielded a main effect of the judgments, $F(1,111) = 9.47, p = .003, \eta^2 = .08$. Independently of the condition, the participants' morality judgments were lower than agency judgments ($M = 5.88, SD = 1.57$ vs. $M = 6.16, SD = .84$). We also found a main effect of the condition, $F(1,111) = 29.12, p = .000, \eta^2 = .21$, and a significant interaction between judgments content and the condition $F(1,111) = 47.23, p = .000, \eta^2 = .30$. A simple effect analysis showed that when the confederate cheated she was perceived as more competent ($M = 5.98, SD = .95$) than moral ($M = 5.03, SD = 1.87$), $t(54) = 5.28, p = .000, d = .64, 95\% \text{ CI } (1.02, 0.25)$. However, when she acted in accordance with norms and did not cheat, she was perceived as less competent ($M = 6.33, SD = .70$) than moral ($M = 6.70, SD = .42$), $t(57) = 4.88, p = .000, d = .64, 95\% \text{ CI } (1.01, 0.26)$. This indicates that the cheating confederate was perceived not only as more competent but also as less moral in comparison to the honest confederate, which strongly suggest that cheating behavior was not perceived as a sign of the lack of specific skills. Previous research showed that people who act for their own interests are perceived as more agentic than those acting against their interests (Cislak & Wojciszke, 2008) and this meshes well with our results and supports the suggestion that selfishly behaving actor has to be competent, even if act of judgment is immoral, because furthering own interests inevitably requires a certain level of agency.

Although the presented study has some limitations, according to our knowledge we showed for the first time that trust in cheaters is not undermined when they act in the observer's interest. This opens new promising avenues for further studies, which could explain more directly why dishonesty leads to trust. According to the dual process theory of moral judgment (Greene, 2007), we can judge actions of other people basing on one of the two systems: automatic (intuitive) or controlled one (deliberative). The presented study suggests that people trust cheaters because they believe in their own moral judgments even when the latter are inflated by self-interest bias. This complements our previous studies (Bocian, Wojciszke, 2014a, b) and suggests the intuitive nature of underlying processes. However, the explanation in terms of deliberative thinking is not entirely ruled out by the present data.

It is risky to trust cheaters with money because they obviously may cheat. Still, people show paradoxical trust in a cheater when the perpetrator cheats and aids their interests. This can help explain why thousands of investors trusted the mega-cheater Bernie Madoff with \$65 billion of their money (McCool & Gryabow, 2009), and these stories continue to repeat in history. Self-interest makes people blind, as La Rochefoucauld noted in the opening quote, and the Madoff-type incidents suggest that even future and uncertain benefits may have this effect.

References

- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*, 596–612. doi: 10.1037//0022-3514.63.4.596
- 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. doi: 10.1037/0022-3514.77.3.525
- Baumard, N. (2010). Has punishment played a role in the evolution of cooperation? *Mind and Society, 9*, 171–192. doi: 10.1007/s11299-010-0079-9
- Bell, R., Schain, C., & Echterhoff, G. (2014). How selfish is memory for cheaters? Evidence for moral and egoistic biases. *Cognition, 132*, 437–442. doi: 10.1016/j.cognition.2014.05.001
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior, 10*, 122–142. doi: 10.1006/game.1995.1027
- Bocian, K., & Wojciszke, B. (2014a). Self-interest bias in moral judgments of others' actions. *Personality and Social Psychology Bulletin, 40*, 898–909. doi: 10.1177/0146167214529800
- Bocian, K., & Wojciszke, B. (2014b). Unawareness of self-interest bias in moral judgments of others' behavior. *Polish Psychological Bulletin, 45(4)*, 411–417. doi: 10.2478/ppb-2014-0050
- Cislak, A., & Wojciszke, B. (2008). Agency and communion are inferred from actions serving interests of self or others. *European Journal of Social Psychology, 38*, 1103–1110. doi: 10.1002/ejsp.554
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology, 104*, 216–235. doi: 10.1037/a0031021
- Cooper, J. (2007). *Cognitive dissonance: 50 years of classical theory*. London: Sage.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. doi: 10.3758/BF03193146
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*, 137–140. doi: 10.1038/415137a
- Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18*, 233–239. doi: 10.1111/j.1467-9280.2007.01882.x
- Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual process theory of moral judgment explains. *Trends in Cognitive Science, 11*, 322–323. doi: 10.1016/j.tics.2007.06.004
- Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition, 106*, 1339–1366. doi: 10.1016/j.cognition.2007.06.007
- Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998–1002. doi: 10.1126/science.1137651
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York, NY: Guilford.
- Levine, E., Schweitzer, M. (2014). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes, 126*, 88–106. dx.doi.org/10.2139/ssrn.2266091
- McCool, G., & Graybow, M. (March 13, 2009). Madoff pleads guilty, is jailed for \$65 billion fraud. *Reuters*. Retrieved in August 1, 2014 from <http://www.reuters.com/article/2009/03/13/us-madoff-idUSTRE52A5JK20090313>
- Rand, D., Greene, J., & Nowak, M. (2012). Spontaneous giving and calculated greed. *Nature, 489*, 427–430. doi: 10.1038/nature11467
- Schulz, J., Fischbacher, U., Thoni, C., Utikal, V. (2014). Affect and fairness: dictator games under cognitive load. *Journal of Economic Psychology, 41*, 77–87. doi: 10.1016/j.joep.2012.08.007
- Stone, J. (2011). Consistency as a basis for behavioral interventions. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 326–347). New York: Guilford Press.
- Suzuki, A., Honma, Y., & Suga, S. (2013). Indelible distrust: Bias towards cheaters revealed as high persistence against extinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1901–1913. doi: 10.1037/a0033335

- Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science, 18*, 689–690. doi: 10.1111/j.1467-9280.2007.01961.x
- Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44*, 1334–1338. doi: 10.1016/j.jesp.2008.03.010
- 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. doi: 10.1037/0022-3514.54.6.1063
- Weisel, O., Shalvi, S. (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences, 112*, 10651–10656. doi: 10.1073/pnas.1423035112

ACKNOWLEDGMENTS

We want to thank Aleksandra Pawłowicz and Olga Zalewska for their help in data gathering. Preparation of the manuscript and data gathering was supported through funding from the Narodowe Centrum Nauki, Poland (Grant NCN 2012/04/A/HS6/00581).