

Personal values and moral motivation 1

Personal values and moral motivation: disentangling moral integrity and moral hypocrisy

Jan-Erik Lönnqvist, Bernd Irlenbusch, Gari Walkowitz

Abstract

Positioning moral motivations within the framework provided by Schwartz' (1992) values theory, we ran three dictator game studies (total $N = 256$) investigating moral integrity and moral hypocrisy. We adapted Batson's (et al., 1997; et al., 1999; et al., 2002) landmark research design into the experimental economics laboratory (Study 1), and showed that the behavioral inconsistency – out of 64 dictators, all 26 who chose to flip a coin to determine the allocation of money ended up with the self-favoring outcome – revealed in such a design is indeed indicative of dishonest claims to morality (arguably the core of moral hypocrisy), and not overpowered moral integrity. Supporting this interpretation, dictators who masked their selfishness behind the coin flip were motivated by high Conformity values (Study 1), and thereby similar to participants who made more obviously disingenuous claims to morality (Study 2). Further, dictators did generally not select the coin flip in case the result could not be rigged (only four out of 32 dictators did this; Study 3). Universalism and Benevolence values were predictive of moral integrity (Studies 1 and 3). Morality ratings of behavior generally showed both self-serving and outcome bias.

Keywords: Personal Values, Moral Motivation, Moral Hypocrisy, Moral Integrity, Ethics

Personal values and moral motivation: disentangling moral integrity and moral hypocrisy

Introduction

Morally dubious behaviors by such professionals as bankers and politicians are far too commonplace to either be listed here or to raise eyebrows. Perhaps more interesting are the recently uncovered systematic scientific frauds carried out by field leaders in the area of moral psychology (see, Alberts, 2011; Callaway, 2011; Gross, 2012). Such frauds cast doubt on the classic Socratean reason for moral failure; i.e., that moral transgressions are caused by intellectual failure. These failures are likely not to have occurred due to lack of knowledge about right and wrong, but rather lack of motivation to behave morally. Indeed, a ground breaking series of studies conducted by Batson and colleagues suggested that many people will not want to pay the cost of behaving morally in case they can avoid it without appearing immoral (Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997; Batson, Thompson, & Chen, 2002; Batson, Thompson, Seuferling, Whitney, & Strongman, 1999).

In the research design employed by Batson and colleagues (et al., 1997; 1999; 2002), moral hypocrisy is revealed through behavioral inconsistency, referring, in the present context, and following the conceptual work of Monin and Merrit (2011), to the discrepancy between what one says and what one does. However, Monin and Merrit (2011) went on to argue that such inconsistency may be only one sign that may (or may not) reveal disingenuous claims to moral virtue – the true essence of moral hypocrisy. The crux of the matter is whether, when claiming the intent to behave morally, people believe that they will behave morally, or whether they are being dishonest. The present research was designed to investigate whether behavioral inconsistency, as revealed in a research design mimicking the one Batson employed, actually indicates moral hypocrisy in the sense of untruthful claims to moral virtue, or whether such

Personal values and moral motivation 3

inconsistency is better interpreted as a sign of another type of moral failure; initial but overpowered moral integrity. More specifically, we first attempted to distinguish such moral hypocrisy and moral integrity in terms of the framework offered by Schwartz' (1992) values theory (Study 1). The results revealed that moral hypocrisy was associated with adherence to Conformity values, whereas moral integrity was associated with adherence to Universalism and Benevolence values. We then showed that adherence to Conformity is associated with dishonestly claiming moral virtue for impression management purposes (Study 2). For Study 3, we altered the experimental setup of Study 1 in such a way that it ruled out disingenuous claims to moral virtue: participants could no longer claim one thing and do another. This served to show that the proportion of people who were actually willing to commit themselves to behaving morally when they knew that they would have to honor their word was extremely low.

Moral Hypocrisy

In a highly influential series of studies, Batson and colleagues (et al., 1997, 1999; 2002) differentiated between moral integrity and moral hypocrisy. Moral integrity was defined as the motivation to actually behave morally. Moral hypocrisy, by contrast, was defined as the motivation to appear as being moral, yet, if possible, avoid the cost of actually behaving morally. Moral hypocrites will enact morality not with an eye on producing a good outcome, but in order to appear moral, yet still benefit themselves. In a typical design, Batson (et al., 1997, Study 2) had participants assign tasks to themselves and an unknown other participant. One task was described as fun and rewarding, the other as boring. Participants were given the option of flipping a coin to assist them make the decision, but it was made clear that the coin flip was not in any way required. Typically, around half of the participants decided not to use the coin, and instead directly assigned the desirable task to themselves. They were acting out of self-interest, but were not hypocrites, as they did not try to conceal their self-interest. Of more interest were those participants who decided to use the coin. Dramatically, across studies, around 90% of coin

flippers flipped the better task for themselves. Around half of Batson's participants claimed they would let the coin decide, but the aggregate data clearly indicated that they did not, suggesting that these people were hypocrites.

Monin and Merrit (2011), in their literature review, updated the classification of moral failures, and provided a more precise definition of moral hypocrisy. They suggested that dishonestly adopting the false appearance of goodness or virtue would be a sufficient and necessary criterion for moral hypocrisy – whether one practices what one preaches is not the central issue (as Batson's definition implies), but whether one believes in what one preaches is. Moral hypocrisy can then be defined as any dishonest claims of morality made to satisfy ulterior, self-serving motives.

Making dishonesty a necessary criterion for moral hypocrisy means that inconsistency between claims and behavior, as revealed in Batson's design, may not necessarily be a sign of hypocrisy. Coin flippers could initially have intended to actually behave morally, and only subsequently, in case the outcome of the flip was unsolicited, given in to the temptation to rig or misreport the coin-toss. That coin flippers, despite laudable intentions, ultimately failed morally could be explained, for instance, by weakness of will. More generally, there is a vast literature on self-regulation that provides an abundance of models (e.g., Baumeister, Bratslavsky, Muraven, & Tice, 1998; Fishbach, Zhang, & Koo, 2009; Shu, Gino, & Bazerman, 2011) explaining why individuals do not always follow through on their intentions. The important point for our purposes is not why people are susceptible to such failures, but rather that this type of behavior-intention inconsistency could explain the results reported on by Batson. Only if participants at the outset had ulterior motives – in essence, never intended to actually let the coin decide – then the behavioral inconsistency observed in Batson's research design would be a sign of moral hypocrisy in the sense of dishonest claims to morality.

Personal values and moral motivation 5

Personal Values

One way to distinguish hypocritical behaviors from initially honest ones could be to examine the motivations underlying such behaviors. For this purpose, the framework offered by Schwartz' (1992) values theory could be especially useful. Values can be defined as transsituational goals that serve as guiding principles in the life of a person or group (Schwartz, 1992). According to Schwartz' model, people in most cultures distinguish between at least ten basic values(the number of cultures in which the same ten basic values have been found was recently reported as 77; Schwartz, 2009) differentiated by motivational content: Universalism, Benevolence, Tradition, Conformity, Security, Power, Achievement, Hedonism, Stimulation, and Self-Direction.

According to values theory, Universalism and Benevolence values are moral values because they express the motivation to promote the welfare of others (Schwartz, 1992). The defining goal of Universalism values is the protection of the welfare of people in general and of nature. Those scoring high on Universalism indicate being committed to, e.g., equality, justice, and protection for all people. At face value, universalists should, in terms of moral integrity and moral hypocrisy, be motivated by the desire to actually be moral, and not merely appear moral. Consistent with the idea that universalists may be particularly prone to care about morality, Universalism has been related to (a) moral sensitivity (Myyry&Helkama, 2002), (b) level of moral reasoning, as conceptualized either through Kohlberg's (1984)stages (Helkama, 2004) or through the moral schemes measured by the Defining Issues Test (Rest, 1979; Myyry, Juujärvi&Pesso, 2010), (c) emotions that motivate moral behavior, such as empathy and guilt (Silfver, Helkama, Lönnqvist, &Verkasalo, 2008), and (d) fairer behavior in standard experimental economics games (Lönnqvist, Walkowitz, Wichardt, Lindeman, &Verkasalo, 2009; Lönnqvist, Walkowitz, Verkasalo, &Wichardt, 2011). Based on these finding, we expected

universalists, when faced with a choice between fair and selfish behavior, to be more likely to behave fairly.

Whereas Universalism values are applied to all members of society (at least in morally inclusive societies, such as Germany, the site of the current research, see Schwartz, 2007), Benevolence values provide an internalized motivational base for preserving and enhancing the welfare of those with whom one is in frequent personal contact. Although most critical in this respect are relations within the family and other primary groups, we assume that the ingroup may comprise also fellow students from one's own university. We therefore expected Benevolence values, like Universalism values, to predict fair behavior. Benevolence values have also been related to emotions that motivate moral behavior (Silfver et al., 2008), but associations with moral reasoning are weak (Helkama, 2004; Myyry et al., 2010).

Conformity values are, both by definition and in the perceptions of laypeople (Schwartz, 2007), also morally relevant. People high in Conformity values will strive not to upset or harm others and will want to behave according to social expectations or norms (Schwartz, 1992; for empirical evidence, see Lönnqvist et al., 2009, Study 1). Because of this tendency to rely on external guidelines, conformists' other personal values could be expected not to predict their behaviors very strongly. Indeed, in the domain of moral behavior, for which the strongest external guidelines are likely to exist (Bardi & Schwartz, 2003), the behavior of conformists is only weakly influenced by their personal Universalism values (Lönnqvist et al., 2009, Study 2; see also Lönnqvist, Leikas, Paunonen, Nissinen, & Verkasalo, 2006), an effect mediated by lack of experienced moral obligation (Lönnqvist et al., 2009, Study 3). Consistent with the above behavioral effects of Conformity values, level of moral reasoning, as conceptualized either through Kohlberg's (1984) stages or the moral schemes of the Defining Issues Test (Rest, 1979), is inversely related to Conformity values (Helkama, 2004; Myyry et al., 2010), further

supporting the view that conformists think of morality in terms of maintaining the conventions provided by society. An important implication of this is that conformists, accustomed to tuning their behavior to social norms – and perhaps morally relevant norms in particular, as these are likely to exert the most normative pressure (Bardi&Schwartz, 2003) – may not internalize those norms. The situation would be analogous to the child who, if paid for home-work, is not motivated to learn (Lepper, Greene, & Nisbett, 1973; Batson, Collins, and Powell(2006) used this analogy in the context of moral hypocrites).Conformists may thus feel so strongly obliged by the values and norms that society provides that questions regarding morality seldom if ever arise, as also suggested by some recent results according to which conformist have difficulties in reporting on a single moral dilemma from their own life (Helkama, 2011).Furthermore, of all ten basic values, Conformity clearly shows the highest correlations with scores on the Marlowe-Crowne social desirability scale (Schwartz, Verkasalo, Antonovsky, & Sagiv, 1997), a measure of the extent to which the respondent seeks to impress others with their good character and gain their approval (Crowne& Marlowe, 1960; see also Paulhus, 1984; 2002). The above summarized work suggests that conformists may pay more attention to appearing moral, rather than to issues of moral right and wrong. Taken together, the above results suggest that conformists may, when faced with a decision between fair and selfish behavior, lack a genuine preference for fairness, but still be highly motivated to appear moral. We therefore expected conformists to use the provided excuse – an ostensibly fair coin flip – to justify their otherwise immoral behavior. In sum, Conformity values were expected to be the motivating force behind moral hypocrisy.

Purpose of the Present Research

We ran three studies using a standard experimental economics game – the dictator game – in order to first establish the occurrence of similar moral failures as those Batson and colleagues (et al., 1997, 1999; 2002) have reported on, and to then augment our understanding of the causes of such failures. The dictator game represents the situation that Batson's participants

faced: the dictator decides on how to split a given amount of money, and the other party is passive. The games that were employed in the three studies are schematically presented in Figure 1.

In Study 1, we employed a within-participants design similar to those used by Batson and colleagues (et al., 1997; et al., 1999; et al., 2002) to examine moral failure revealed by inconsistency between what one says and what one does. Dictators could either directly choose the fair or selfish allocation of money, or they could decide by flipping a labeled coin (see Batson et al., 2002, Study 2) that was provided for this purpose. Based on the results of Batson's research program, we expected a substantial number of participants to claim using the coin, but to reveal, through their aggregate behavior, that they did not abide by the result of the coin flip. Regarding personal values, we expected participants high in Universalism or Benevolence values to directly choose the fair allocation of money. However, among participants who chose the selfish allocation, we expected those high in Conformity to be more likely to claim to have flipped the coin to arrive at the selfish outcome.

Previous research on social dilemmas has established that individuals' make more moral decisions if they are watched by others (e.g., Bateson, Nettle, & Roberts, 2006; Keller & Pfattheicher, 2011; Wedekind&Milinski, 2000). To investigate this, we included a manipulation designed to heighten the dictators' sense of being watched. We expected choosing the coin to cover up selfish motives to increase in the condition of being watched. Furthermore, we expected conformists to be especially sensitive to this manipulation (Keller & Pfattheicher, 2011; Leikas, Lönnqvist, Verkasalo, & Lindeman, 2009).

We also gathered ratings regarding the morality of the various behaviors. Directly choosing the fair (selfish) option was expected to be rated as the most (least) moral course of action. Based on a recent study that suggested that responsibility can to some extent be shifted

even to a random mechanism (Bartling&Fischbacher, 2012), we expected the coin flip to be rated somewhere in between directly choosing either the fair or selfish option. Furthermore, evaluations were expected to be outcome dependent (Gino, Moore, & Bazerman, 2009): the flip was expected to be rated as a more moral course of action if it led to the fair rather than selfish decision. Finally, we investigated whether the evaluations of morality were self-serving. Based on the results of Batson and colleagues (et al., 1997; et al., 1999, et al., 2002), we expected coin flippers to rate this course of action as more moral than those who did not flip the coin, revealing moral hypocrisy in the sense applying moral standards for self-serving purposes (Monin&Merrit, 2011).

In Study 2, drawing on the framework suggested by Monin and Merrit (2011), we expected a between-participants manipulation of monetary incentive to reveal dishonest claims to morality. Participants asked how they would split the money were expected to declare fairer allocations, as compared to the behavior of those participants who actually allocated the money (the option to let the coin decide was not given in Study 2). More important, we expected conformists to be especially likely to adjust their behavior as a function of whether or not real money was at stake. This would testify to the tendency of conformists to make dishonest claims to morality in the service of impression management motives, suggesting that those who flip the coin in Study 1, if driven by Conformity values, are indeed dishonest from the very beginning. Regarding the other moral values, we expected, across conditions, those high in Universalism or Benevolence to be more likely to select the fair option.

To further examine whether participants in Study 1 were initially intending to actually let the coin decide, we introduced in Study 3 the binding coin. Would participants, at the outset, choose to flip a coin that could not be rigged, thereby indicating an initial motivation to behave

morally? Regarding personal values, we again expected Universalism and Benevolence values to predict choosing the fair option at the outset.

Study 1

Study 1 represents Batson's research design in the form of a dictator game. Our purpose was to connect moral integrity and moral hypocrisy – as revealed through within-participants behavioral inconsistency – to personal values.

Methods

Participants and Procedure. Study 1 was conducted with 128 participants (mean age=24.3 years ($SD=2.87$); 52.3% female) from the University of Cologne (Germany) majoring in different disciplines. Participants were invited through ORSEE (Greiner, 2004). Four experimental sessions were run involving 64 participants per treatment and 32 participants per session. On showing up, participants were randomly seated in computer cubicles that secured anonymity, randomly assigned a role (dictator or passive other participant), and randomly matched in dyads of two participants. All experimental sessions were conducted on the computer using z-Tree (Fischbacher, 2007).

Measures. In the dictator game two participants are randomly matched and assigned either the role of the dictator or the passive other person (referred to as Person A and Person B in the instructions). At the outset of the game, the dictator is provided with an amount of 10€ which is to be distributed among herself and the other person. All dictators were additionally provided with a coin that was labeled “8/2” on one side and “5/5” on the other side (Batson et al., 2002, Study 2). The dictator could choose from three options: 1) she could either directly choose the option “8/2” that yielded 8€ for herself and 2€ for the other person; 2) she could directly choose the option “5/5” that yielded equal payoffs of 5€ for both; or 3) she could flip the coin and determine the distribution according to the outcome of the coin flip. The dictator was asked to

indicate her choice (“8/2” or “5/5”), and whether or not she had flipped the coin (“Yes” or “No”). The other person did not make any decisions (see Figure 1, panel A). Furthermore, we included a manipulation designed to induce the feeling of being watched. Before acting, half of the dictators were informed that the other person would, via the computer network, be notified of whether or not the dictator indicated having used the coin (watched condition). In the other condition, the other person was not notified about coin use (unwatched condition).

Regarding evaluations of morality, both dictators and receivers rated the morality (0 = *not at all moral*; 8 = *very moral*) of dictators’ four possible courses of action; i.e., the morality of choosing (a) 5/5 without coin, (b) 5/5 with coin, (c) 8/2 without coin, and (d) 8/2 with coin. These evaluations were gathered after the dictator had made her decision.

The Portraits Values Questionnaire (Schwartz et al., 2001) was, across studies, administered at the end of the session to measure the ten basic values. This 40-item questionnaire consists of descriptions of a hypothetical person in terms of his or her goals and aspirations. Respondents are to rate their similarity to the person on each item, using a 6-point scale ranging from 1 (*not like me at all*) to 6 (*very much like me*). The measure includes six items for measuring Universalism (example item: She wants everyone to be treated justly, even people she doesn’t know. It is important to her to protect the weak in society), four for Benevolence (example item: It’s very important to her to help the people around her. She wants to care for their well-being), and four for Conformity (example item: It is important to her always to behave properly. She wants to avoid doing anything people would say is wrong). In computing the values scores, we used proportional sum variables that centered, within participants, the average of the ten basic values at one (Schwartz, 1992). The descriptive statistics were: Universalism ($M = 1.04$, $SD = 0.18$, $\alpha = .77$), Benevolence ($M = 1.14$, $SD = 0.16$, $\alpha = .62$), and Conformity ($M = 0.89$, $SD = 0.18$, $\alpha = .64$).

Results

Behavioral results. Sex, age, and condition had no effects on any of the dependent variables, and were collapsed for correlational analyses. However, any statistically significant correlations were further investigated in regression analyses that controlled for both the condition (watched vs. unwatched) and the interaction between the condition and the investigated personal value (condition x value interaction), as well as for age and sex. Of the 64 participants who were assigned the role of the dictator, 56 chose the selfish option, of whom 26 (46%) indicated having flipped the coin (behavioral decisions across the three studies are shown in Table 1). Out of the 26 participants who indicated having flipped the coin, all reported having flipped the selfish option (8/2). The likelihood that a fairly flipped coin would give such an outcome is extremely small ($p < .0000001$), suggesting that participants were not being truthful.

As expected, both Universalism ($r(63) = .29, p = .02$) and Benevolence ($r(63) = .34, p = .01$) were correlated with choosing the fair option at the outset (choosing 5/5 was coded as 1, choosing the selfish outcome, with or without coin flip, was coded as 0). Conformity values were not correlated with choosing the fair option ($r(63) = .06, p > .10$). In a regression analysis that predicted choosing the fair option, and with condition (watched vs. unwatched), age, sex, Universalism, and the interaction between condition and Universalism values entered as predictor variables, the main effect of Universalism values was statistically significant ($\text{Beta} = .45, t(58) = 2.48, p = .02$). In a similar regression analysis involving Benevolence values, the influence of Benevolence was rendered marginally statistically significant ($\text{Beta} = .32, t(58) = 1.80, p = .08$). No other variables emerged as statistically significant predictors in either regression analysis.

Among the 56 participants who chose the selfish outcome, Conformity values were, as expected, correlated with claiming to have used the coin to obtain this outcome ($r(63) = .27, p = .05$). Conformity remained a statistically significant predictor ($\text{Beta} = 0.42, t(50) = 2.18, p = .03$)

in a regression analysis that included, besides Conformity, condition, age, sex, and the interaction between condition and Conformity.

Evaluations of morality. Condition (watched vs. unwatched), age, and sex had virtually no effects on morality ratings; participants were again collapsed for correlational analyses. We had four different groups from which we obtained ratings of the morality of the various decisions: (1) passive other persons ($N = 64$); (2) those dictators ($N = 8$) who directly chose the fair option; (3) those dictators ($N = 30$) who directly chose the selfish option; and those dictators ($N = 26$) who claimed to have flipped the coin to arrive at the selfish option. The means of these four groups' morality ratings are shown in Table 2. Inspection of the means shows that choosing the fair option without the coin flip was, as expected, perceived as the most moral course of action. This choice was, across all participants, perceived as more moral than choosing the fair option as a result of the coin flip ($d = 0.98$, $t(127) = 9.06$, $p < .001$; the t-values are from paired samples t-tests, but the effects sizes are computed from the original means and standard deviations (Dunlop, Cortina, Vaslow, & Burke, 1996)). Also as expected, choosing the selfish outcome was, across all participants, perceived as more moral if preceded by the coin flip than if chosen directly ($d = 0.97$, $t(127) = 9.06$, $p < .001$).

Post-hoc tests comparing the above four groups showed they were quite similar in their moral evaluations, with the exception of those participants who chose the fair outcome (Table 2). These participants could be described as applying stricter moral standards. To further examine group differences, we computed, as an index of the power of the coin to render the selfish outcome more moral, the within-participants difference score between morality ratings of the selfish outcome with and without the coin flip. This difference score is referred to as the Coin Effect (shown on last row of Table 2). Examination of the means across groups shows that those participants who indicated having flipped the coin in order to obtain the selfish outcome rated this course of action as more moral than did those who were passive or those who directly chose

the selfish outcome. The latter two groups did not differ from each other, making, as expected, the moral evaluations of the coin flippers stand out, thereby highlighting the self-serving nature of these evaluations.

Discussion

The results of Study 1 provide firm evidence of moral failure as revealed through behavioral inconsistency. Replicating the results of Batson and colleagues (et al., 1997; et al., 1999; et al., 2002), all 26 participants who indicated having determined the distribution of money by flipping the coin ended up with the selfish outcome. Both Universalism and Benevolence values were correlated with choosing the fair option, suggesting that these values can help explain moral integrity. By contrast, among those participants who chose the selfish option, those high in Conformity were more likely to use the coin flip as an excuse for their selfish behavior. Considering previous work on Conformity (see Introduction), this result suggests that impression management processes at least partly motivated the choice of the coin flip.

Choosing the selfish option was regarded as more moral if preceded by a coin flip. Even in a context in which a fair alternative option is present, the coin flip cannot thus be considered merely a lame excuse for behaving immorally (cf. Batson et al., 2003). By contrast, such a random mechanism appears to provide a plausible excuse for selfish behavior (see also Bartling&Fischbacher, 2012).Even passive participants, and those who directly chose the fair option, considered the selfish option as somewhat more moral if preceded by a coin flip. However, the effect of the coin to render the selfish option more moral was particularly strong among dictators who indicated having flipped the coin. This suggests that such dictators may not be applying ethical standards objectively, but rather for other motives, making such evaluations hypocritical (Monin&Merrit, 2011). Besides showing such self-serving bias, the moral evaluations also showed clear outcome bias (Gino, Moore, &Bazerman, 2009): using the coin

and arriving at the fair outcome was evaluated as more moral than flipping the coin and obtaining the selfish outcome. However, flipping the coin and thereby obtaining the fair option was not perceived to be as moral as directly choosing the fair option. This pattern resembles the results of Bartling and Fischbacher (2012). They found that when the dictator shifted responsibility to a random mechanism, the dictator was strongly punished when the outcome was unfair. However, even a fair outcome could result in punishment if the outcome was obtained through the random mechanism.

In contrast to Batson's design (et al.; 1997; et al.; 1999; et al.; 2002), dictators could make a fair choice at the outset: the choice of the coin was, to some degree, already indicative of moral failure. But were those participants who flipped the coin intending to abide by the result, thereby giving the other person a fair chance (indicating moral integrity – note that this behavior was generally rated as more moral than just choosing the selfish outcomes)? Or were they flipping the coin merely in order to appear moral? The fact that flipping the coin was associated with Conformity values would suggest the latter option. We further investigate these questions in Studies 2 and 3.

Study 2

Study 1 showed clear evidence of inconsistency between claims and behavior. But were those who flipped the coin intending to abide by the results of the coin flip, or were they doing so merely to appear moral? As such behavior was most typical for those scoring high in Conformity values, the latter option might be more likely. But can Conformity values really be thought of as indicative of the disposition to make dishonest claims to morality? In Study 2, we used a between-participants manipulation to expose dishonest claims to morality.

Methods

Participants and Procedure. Study 2 was run with 64 participants (two sessions with 32 participants; average age=24.2, $SD=2.61$; 51.6% female) from the University of Cologne. The

procedure was similar to that of Study 1.

Measures. The dictator game in Study 2 was similar to the one in Study 1, except that there was no coin, and both players made a choice. After receiving an endowment of 10€, the dictator made a monetarily incentivized decision between the “8/2” or the “5/5” distribution (incentivized condition; see Figure 1, panel B), whereas the other player made the corresponding hypothetical choice (hypothetical condition).

Personal values were measured by the Portraits Values Questionnaire (Schwartz et al., 2001). The descriptive statistics were: Universalism ($M = 1.08$, $SD = 0.23$, $\alpha = .84$), Benevolence ($M = 1.11$, $SD = 0.21$, $\alpha = .76$), and Conformity ($M = 0.89$, $SD = 0.21$, $\alpha = .74$).

Results

Age and sex had no effects on any of the outcomes. Regarding the effect of condition, of the 32 participants who made an incentivized choice, 26 chose the option that favored themselves, whereas only six chose the fair option. By contrast, in the hypothetical condition, the majority of participants, 17 out of 32, chose the fair option. As expected, the fair option was chosen more often in the hypothetical than in the incentivized condition ($\chi^2(1) = 6.78$, $p = .01$).

To examine the effects of Universalism, Benevolence, and Conformity, we ran three separate regression analyses, one for each basic personal value, predicting choice behavior with condition (incentivized vs. hypothetical), age, sex, one of the three basic values, and the interaction between that value and condition. The expected interaction between condition and Conformity values was statistically significant ($Beta = 0.99$, $t(58) = 2.00$, $p = .05$). This interaction is plotted in Figure 2. As expected, in the hypothetical condition, with age and sex also entered as predictor variables, Conformity values predicted choosing the fair option ($Beta = 0.32$, $t(28) = 2.01$, $p = .05$). However, this was not true in the incentivized condition ($Beta = -.08$, $t(28) = -.42$, $p > .10$). Neither Universalism nor Benevolence showed the expected main

effect on behavior, and neither value interacted with the condition.

Discussion

The results of Study 2 showed, as expected, that many participants were motivated to convey, without cost, a moral impression. However, very few were willing to pay the price for actually behaving morally. In essence, most participants in the hypothetical condition claimed that they would choose the fair rather than selfish option. However, the disingenuous nature of this claim (i.e., that participants were distorting their responses in order to appear more moral, either to the experimenter or some generalized imagined other; see Baumeister, Tice, & Hutton, 1989) was revealed in the monetarily incentivized condition, in which participants were more often selfish.

As expected, particularly those scoring high in Conformity values made disingenuous claims to morality. Such hypocrisy suggests that the behaviorally inconsistent participants in Study 1, who were also motivated by Conformity, were from the outset choosing to flip the coin only for the sake of appearing moral, with little intention of actually behaving morally. In Study 3 we sought further evidence that it is, indeed, impression management from the very beginning that is at the core of the moral failure observed in Study 1. The lack of findings for Universalism and Benevolence is most likely best explained by the relatively small sample size of Study 2.

Study 3

Study 2 showed that the same Conformity values that were associated with Study 1 participants' behavioral inconsistency were also associated with, in our view, disingenuous claims to morality. However, we acknowledge that it is possible to doubt whether the claims made in the hypothetical condition in Study 2 actually were disingenuous. Perhaps participants actually believed that they would have made the fair choice had there been money at stake. This would mean that neither Study 1 nor Study 2 would definitely showcase moral hypocrisy in the sense of "preaching in bad faith" (Monin & Merrit, 2011), but rather demonstrate that conformists

will show a particularly strong discrepancy between their beliefs regarding their behavior and how they actually behave. This is not our interpretation of the results, as impression management concerns appear to be much closer to the core of Conformity values. But as the possibility for such an interpretation exists, we further investigated the likelihood that Study 1 participants were initially intending to let the coin decide. For this purpose, we introduced in Study 3 the binding coin.

Methods

Participants and Procedure. Study 3 was executed with 64 participants (32 participants per session; average age=24.5, $SD=3.1$; 56.3% female) from the University of Cologne (Germany). The procedure was similar to that of Study 1.

Measures. We changed our experimental setup from Study 1 by enforcing the coin flip. In essence, participants who chose the coin flip to determine the outcome were forced to abide by the result of coin flip. At the outset of the game, the dictator again received an endowment of 10€ which was to be distributed among herself and the passive other person. As in Study 1, the dictator could 1) directly choose the option “8/2”, 2) directly choose “5/5”, or 3) she could shift her decision to a fair lottery implemented by the experimenters. This lottery (which was run on the participant’s computer) was structurally identical to the coin in Study 1 in that it yielded with a probability of 50% the option “8/2” and with a probability of 50% the option “5/5”. The other player did not take any action (see Figure 1, panel C).

Personal values were measured with the PVQ (Schwartz et al., 2001). The descriptive statistics were: Universalism ($M = 1.03$, $SD = 0.15$, $\alpha = .75$), Benevolence ($M = 1.16$, $SD = 0.14$, $\alpha = .64$), and Conformity ($M = 0.88$, $SD = 0.20$, $\alpha = .53$).

Results

Age and sex had virtually no effects on any of the outcomes. Out of the 32 decision-

makers, six chose the fair option at the outset, 22 chose the selfish option, and four (12.5%) chose the binding coin flip. Compared to Study 1, in which 26 participants out of 64 (40.6%) flipped the coin, the number of participants who now chose the coin was significantly smaller (comparing across studies, $\chi^2(1) = 6.60, p = .01$).

Universalism values predicted choosing the fair option at the outset ($r(31) = .37, p = .04$; directly choosing the fair option was coded as 1, other choices as 0). Universalism retained its predictive power in a regression that also included age and sex as predictor variables (Beta = .41, $t(28) = 2.29, p = .03$). A similar, although only marginally statistically significant association was found between Benevolence values and choosing the fair option ($r(31) = .33, p = .07$). This association was also rendered statistically significant when controlling for age and sex (Beta = .41, $t(28) = 2.17, p = .03$).

Discussion

Comparing the results of Study 3 with those of Study 1 shows that although half of the participants claimed to have abided by the result of a coin flip (Study 1), very few were actually willing to commit themselves to a binding coin (Study 3). This means that the decision in Study 1 to flip the coin was most likely not determined by the motivation to behave morally, but rather by the motivation to appear moral. In essence, coin flippers in Study 1 did not at the outset have good intentions that they were unable to follow through on due to lack of willpower. By contrast, such participants were most likely not planning to live up to the moral ideal, but were merely concerned with their moral appearance.

Regarding personal values, the number of participants who chose the binding coin flip was rather small, not allowing for conclusions regarding the values of those people who made this choice. However, further supporting the results of Study 1, those high in Universalism or Benevolence values were, as expected, more likely to choose the fair option at the outset.

General Discussion

Considering the lack of credibility of many public figures that currently plagues an increasing number of social psychologists (see opening paragraph), it is reassuring that the results reported on in one of the most important series of studies in moral psychology replicated in another laboratory, in another country, and with non-overlapping researchers. This happened although our design was in many ways very different from that used by Batson and colleagues, suggesting that the moral failure described by Batson is robust across a wide range of contextual variables. Even more important, our results provide evidence that the behavioral inconsistency observed in Batson's research design is actually a sign of moral hypocrisy, in the sense of disingenuous claims to morality (Monin & Merrit, 2011), and not overpowered moral integrity. The other major contribution of the present research was to show that individual differences – conceptualized within the framework offered by Schwartz' values theory – could help explain and disentangle both moral integrity and moral hypocrisy.

Behavioral Inconsistency and Disingenuousness Claims to Morality

As noted by Monin and Merrit (2011), not all instances of behavioral inconsistency are necessarily instances of moral hypocrisy, assuming the latter is defined as disingenuous claims to morality. Indeed, participants in Study 1 who chose to flip the coin – around half of the participants – could at the outset have been motivated to behave at least somewhat morally (the truly moral thing to do would of course have been to choose the fair option). Their failure to follow through on their moral intentions, as revealed by the fact that all of them claimed to have flipped the selfish outcome, could thus have been an instance of overpowered moral integrity rather than moral hypocrisy.

Study 1 showed that participants scoring high in Conformity were particularly likely to claim having flipped the coin. Considering that those scoring high in Conformity may think it important to appear moral (see Introduction), the result suggests the importance of impression

management motives rather than overpowered moral motives. In Study 2 we investigated whether Conformity could be related to a more straightforward instance of dishonesty. The results of Study 2 first revealed that people were generally willing to portray themselves as moral when this could be done without cost. More importantly, particularly those scoring high in Conformity values were prone to claim that they would behave morally if it was their task to allocate the money. The dishonesty of this claim was revealed in the monetarily incentivized condition, in which conformists were not more likely to behave morally. In Study 3 we sought even stronger evidence that it is, indeed, merely the motivation to appear moral that is at the core of the moral failure observed in Study 1. Indeed, when faced with a binding coin flip, participants were generally not willing to commit themselves to the coin, but preferred directly choosing the selfish option (this is consistent with the results of Bartling & Fischbacher, 2012). Together these results suggest that the moral failure revealed in Batson's research design through behavioral inconsistency can actually be categorized as moral hypocrisy in the sense of making disingenuous claims to morality.

As the behavioral outcome does not vary, one could ask whether the motivation behind moral failures really matters. Besides advancing our theoretical understanding of a very subtle and complex issue that has attracted the attention of philosophers for centuries – why do good people behave immorally – our results could also have applied relevance. If moral hypocrisy, as our results suggest, does not result from weakness of will, then institutional changes that are known to heighten self-control (e.g., motivational incentives, training on self-control tasks, and glucose supplementation; for a review, see Hagger, Wood, Stiff, & Chatzisarantis, 2010) are not likely to be very successful in deterring moral failures. Indeed a recent review and meta-analysis on business ethics instructional programs suggests that they have at best minimal impact (Waples, Antes, Murphy, Connelly, & Mumford, 2009). Instead, if many people merely care about appearing moral, changes directed at increased transparency and exposure should be much

more effective deterrents. The exact nature of such institutions would of course vary from one field to another (e.g., depositing the raw data and computer codes into public data archives, as required by some scientific journals, should increase transparency (e.g., McCullough, McGeary, & Harrison, 2006), whereas the ‘Clean Hands’ campaign in Italy (Della Porta, 2001) was intended to increase the risk of being caught). It is also noteworthy that some of the participants were actually motivated to be moral – especially those scoring high in Universalism values. Unfortunately, universalists may be less likely to be drawn into fields such as business (e.g., Lindeman & Verkasalo, 2005; Lönnqvist, Verkasalo, Helkama, Andreyeva, Bezmenova, et al., 2007), suggesting that they cannot be looked upon to halt the continuation of large-scale moral failures, at least inside the business world.

Moral Values

Batson’s research program assessed individual differences using questionnaire measures of dispositions commonly thought to be associated with moral responsibility. However, his results showed that these questionnaires primarily measured the motivation to appear moral in the eyes of oneself and others, not the actual propensity to behave morally. One important difference, discussed below in more detail, between Batson’s research design and ours was that Batson’s design lacked an outcome fair option. This may be one of the reasons that Batson and colleagues found various individual difference measures of moral responsibility to predict primarily moral hypocrisy and not moral integrity – even morally responsible participants may have found it unreasonable to assign or even risk assigning the other participant to the positive task. By contrast, in the presence of an outcome fair option, morally responsible participants should have far less excuses to do anything but pick this option. This option may also have functioned to activate Universalism and Benevolence values: even personally important values are known to predict behavior only when activated (Verplanken & Holland, 2002). In fact,

considering the theoretical dispute (for a review, see Maio, Olson, Bernard, Luke, 2003) and empirical lack of research on the possible association between personal values and behavior (e.g., of the illustrations of value-behavior relations reviewed by Roccas and Sagiv (2010), Lönnqvist, Verkasalo, Wichardt, and Walkowitz (2012) argued that none involved actual behavior), one of the major contributions of the present research is that it shows the importance of personal values in determining (moral) behavior.

The present results testify to the importance of assessing values at the level of the ten basic values. Although Benevolence and Conformity values are adjacent to each other in the values circle (Schwartz, 1992), their implications for moral behavior are very different. Such difference would most likely not be captured by broader measures, such as measures of the two broad value dimensions that organize the values circle (e.g., Verkasalo, Lönnqvist, Lipsanen, & Helkama, 2009). As in personality psychology more generally (Paunonen & Ashton, 2001), the predictive power of narrower constructs appears superior to that of broader constructs. Rather than exclusively focus on broader and shorter measures, as has been the recent trend in personality and social psychology (e.g., Denissen, Geenen, Selfhout, & van Aken, 2008; Lindeman & Verkasalo, 2005; Rammstedt & John, 2007), our results suggest that it might for many purposes be fruitful to measure values at an even more specific level than that of the ten basic values.

Differences between our and Batson's Research Design

Our decision to adapt Batson's design rather than directly replicate it was in part motivated by some recent criticism suggesting that Batson's results on moral hypocrisy are, in fact, a product of experimenter demand effects. Specifically, making the norms of justified self-interest, morality, and fairness salient, has been shown to cause experimenter demand effects that could account for Batson's findings (Fernández-Dols et al., 2010). We were careful not to

elicit such demand effects in the present design. Furthermore, as the present research was run in an experimental economics laboratory, the ethical guidelines followed by this community also forced us to adjust the research design. These guidelines posit, for instance, that no deception of participants is allowed.

Fernández-Dols and colleagues (2010), after providing evidence that the hypocrisy revealed by Batson's research could be a phenomenon produced by experimenter demand effects, went on to suggest that in real life, hypocrisy, because of its social dysfunctionality and the derogation of hypocrites, is infrequent. Therefore, although our experimental design was different than Batson's, comparison between the number of participants who displayed moral hypocrisy may still be interesting. In Study 1, 40% ($N = 26$) of our participants flipped the coin, and many of them were with high probability hypocrites, as revealed by the fact that a 100% flipped the selfish option for themselves. These numbers are highly similar to the ones reported by Batson and colleagues. In the setup closest to our research design, Batson et al. (1997; study 2) reported that 10 out of 20 (50%) participants flipped the coin, and nine of those ten (90%) assigned themselves the fun task. Thus, although our research design did not involve the type of experimenter demand effects identified by Fernández-Dols and colleagues (2010) as possibly causing the moral hypocrisy phenomenon, it pointed to the high frequency of moral hypocrisy.

One of the most important differences to earlier studies on moral hypocrisy was that we gave participants the opportunity to be fair regarding potential outcomes. The introduction of a fair option meant that choosing the coin, regardless of the result of the flip, was already, by comparison, a moral failure (as was correctly reflected in the morality ratings). In the presence of an outcome-fair (or distributively fair) option, the excuse provided by a procedurally fair option (the random coin flip) has previously been suggested to be too lame an excuse to appeal to those seeking to cover up their selfish motivation (Batson et al., 2003). However, almost half of our

participants did try to hide their selfish motive behind the coin flip. Perhaps outweighing the ostensible lameness of the excuse was the steeper threshold of directly choosing the selfish option: self-serving decisions are harder to make in the presence of a fair option (Scheepers, Spears, Manstead, & Doosje, 2009). The introduction of an outcome-fair option could thus have had two effects that worked in opposite directions: although such an option makes the excuse offered by the process-fair option (coin flip) look lame, it also makes it more difficult to directly choose the selfish option. Interestingly, participants did not react to the potential unattractiveness of these two options by choosing the third, outcome-fair option (only 13 % of Study 1 participants did this). Instead, testifying to the pervasiveness of lack of moral integrity, nearly all participants continued to choose the selfish outcome, and around half of them excused their behavior with the coin.

When considering generalizability out of the laboratory context, the presence of a fair outcome may serve to increase the external validity of the research: in many real-life conflicts of interest, an outcome that is mutually perceived as fair is often negotiable. From this perspective, another important difference to the research design employed by Batson and colleagues was that our participants were not assigning tasks, but splitting a real amount of money. Money could increase the external validity of the research by connecting laboratory decisions to the world outside of the laboratory (Falk & Heckman, 2009; Levitt, & List, 2007). In fact, the results of Study 2 support the idea that whether or not behavior is monetarily incentivized to some extent determines behavior: although the majority of people claimed that they would behave morally in circumstances in which such claims were costless, less than a fifth actually behaved morally when the behavior had monetary consequences (see also Lönnqvist, Verkasalo, & Walkowitz, 2011).

Another important advantage of having participants make monetary decisions is that there is no need for deception. The gullibility of the research participants therefore does not

become an issue, and there is no need to retrospectively determine whether participants actually believed in the cover story (e.g., that there was another participant involved). Most importantly, no participants have to be excluded from the data set due to suspicion regarding the research design. This eliminates one source of potentially systematic error that could distort the results.

Moral Evaluations

In general, the decision to use the coin, thereby giving the fair option an evenhanded chance, was rated as more moral than the decision to directly choose the selfish option. Consistent with the results reported on by Bartling and Fischbacher (2012) some of the blame could thus be shifted to a chance mechanism, despite the fact that dictators had the option to directly choose the fair option. Regarding the self-serving nature of moral evaluations, our results confirm the results reported on by Batson and colleagues (et al., 1997; et al., 1999; et al., 2002) in the sense that especially those participants who flipped the coin (and obtained the selfish outcome) perceived this as a more moral course of action than did those who did not flip the coin. Extending upon Batson's results, we could compare the morality ratings of the dictators with those of passive other players. Without such independent ratings, it would not be possible to distinguish whether it is the coin flippers or the other dictators who are self-serving in their ratings of morality. For instance, participants who directly chose the selfish option could themselves be hypocrites in denying the virtue of behavior that could in comparison threaten their self-image (Monin, Sawyer, & Marquez, 2008). Or, participants who directly chose the fair option could out of jealousy condemn the coin flippers (for the "sucker-to-saint" compensation effect, see Jordan & Monin, 2008), who managed to obtain the selfish outcome whilst still appearing moral. The other, passive, participants' ratings of morality could be considered independent in the sense that they had not made any choice. Comparing their outside perspective with that of the dictators', it is clear that it is the coin flippers who deviate from such an

independent evaluation of morality: coin flippers perceived choosing the selfish outcome as a result of the coin flip, as compared to directly choosing it, as more moral than those who did not make any choice. This could be a further sign of moral hypocrisy, in the sense that coin flippers were not objective in their application of moral standards (Valdesolo & DeSteno, 2008; Monin & Merrit, 2011). However, it could also be that coin flippers chose to flip the coin because they perceived this course of action as more moral than did other participants.

The moral evaluations were also outcome dependent: the coin flip was rated as more moral if it led to the fair rather than selfish outcome. This is consistent with the results of Gino et al. (2009), who found outcome dependence in ethical judgments of actions presented in short vignettes. The current design differs from that reported on by Gino et al. (2009) in that the outcomes had differing monetary consequences for our participants. However, whether or not one had decided to use the random mechanism, or whether or not the outcome was favorable to oneself, the outcome bias persisted in ethical judgments (there were no differences in outcome bias between the groups reported in Table 2, results not shown), suggesting the pensiveness of the bias.

Conclusions

Perhaps the central contribution of Batson's research on moral hypocrisy is that it reveals that moral failures are due to a lack of moral motivation; i.e., people are primarily motivated to appear moral, and not to be moral. Batson and colleagues have investigated and rejected alternative accounts of their results, such as misperception of hypocritical behavior as moral (Batson et al., 1999; Study 1) and low salience of social standards (Batson et al., 2002; Study 1). We add to this literature by showing that the behavioral inconsistency revealed in Batson's research design indeed indicates disingenuous claims to morality –the core of moral hypocrisy (Monin&Merrit, 2011) – and not an initial attempt to be moral that only later is foiled (overpowered moral integrity).

The other important contribution of the present research is that we position both moral integrity and hypocrisy into the framework provided by Schwartz' values theory. Whereas the former is motivated by Universalism and Benevolence values, the latter is driven by adherence to Conformity values. Connecting moral integrity and moral hypocrisy to established and well-researched constructs such as personal values should illuminate promising avenues for future research.

References

- Alberts, B. (2011). Expression of Concern on Stapel and Lindenberg. *Science*, 332, 251-253.
- Bardi, A. B., & Schwartz, S. H. (2003). Values and behaviour: Strength and structure of relations. *Personality and Social Psychology Bulletin*, 29, 1207–1220.
- Bartling, B., & Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *Review of Economic Studies*, 79, 67-87.
- 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., Collins, E. C., & Powell, A. A. (2006). Doing business after the fall: The virtue of moral hypocrisy. *Journal of Business Ethics*, 66, 321-335.
- Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, D. W. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, 72, 1335–1348.
- Batson, C. D., Lishner, D. A., Carpenter, A., Dulin, L., Harjusola-Webb, S., Stocks, E. L., et al. (2003). “... As you would have them do unto you”: Does imagining yourself in the other’s place stimulate moral action? *Personality and Social Psychology Bulletin*, 29, 1190-1201.
- 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., Seuferling, G., Whitney, H., & Strongman, J. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, 77, 525–537.
- Baumeister, R. F., Tice, D. M., & Hutton, D. G. (1989). Self-presentational motivations

- and personality differences in self-esteem. *Journal of Personality*, 57, 547-579.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.
- Callaway, E. (2011). Report finds massive fraud at Dutch universities. *Nature*, 479, 15.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Della Porta, D. (2001). A judges' revolution? Political corruption and the judiciary in Italy. *European Journal of Political Research*, 39, 1-21.
- Denissen, J. J. A., Geenen, R., Selfhout, M., & van Aken, M. A. G. (2008). Single-item Big Five ratings in a social network design. *European Journal of Personality*, 22, 37–54.
- Dunlop, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170–177.
- Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326, 535 - 538.
- Fernández-Dols, J.M., Aguilar, P., Campo, S., Vallacher, R., Janowsky, A., Rabbia, et al. (2010). Hypocrites or maligned cooperative participants? Experimenter induced normative conflict in zero-sum situations. *Journal of Experimental Social Psychology*, 46, 525-530.
- Fishbach, A., Zhang, Y., & Koo, M. (2009). The dynamics of self-regulation. *European Review of Social Psychology*, 20, 15–344.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, 171-178.

- Gino, F., Moore, D. A., & Bazerman, M. H. (2009). No harm, no foul: The outcome bias in ethical judgments. Unpublished manuscript.
- Greiner, B. (2004). The Online Recruitment System ORSEE 2.0 - A guide for the organization of experiments in economics. *University of Cologne Working Papers, 10*.
- Gross, C. (2012, January 9-16). Disgrace: On Marc Hauser. *The Nation*.
- Hagger, M.S., Wood, C., Stiff, C., & Chatzisarantis, N.L.D. (2010). Ego depletion and the strength model of self-control: A metaanalysis. *Psychological Bulletin, 136*, 495–525.
- Helkama, K. (2004) Values, role-taking and empathy in moral development. *New Review of Social Psychology, 3*, 103-111.
- Helkama, K. (2011). Social psychology of morality and moral development: A functional system approach. In J. P. Valentim (Ed.), *Societal approaches in social psychology* (pp. 131-149). Frankfurt am Main: Peter Lang GmbH Company.
- Kohlberg, L. (1984). *The psychology of moral development*. San Francisco, Harper & Row.
- Jordan, A. H., & Monin, B. (2008). From sucker to saint: Moralization in response to selfthreat. *Psychological Science, 19*, 809-815.
- Keller, J. and Pfattheicher, S. (2011). Vigilant self-regulation, cues of being watched and cooperativeness. *European Journal of Personality, 25*, 363–372.
- Leikas, S., Lönnqvist, J.-E., Verkasalo, M., & Lindeman, M. (2009). Regulatory focus systems and personal values. *European Journal of Social Psychology, 39*, 415-429.

- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the overjustification hypothesis. *Journal of Personality and Social Psychology, 28*, 129–137.
- Levitt, S. D. & List, J. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives, 21*, 153-174.
- Lindeman, M., & Verkasalo, M. (2005). Measuring values with the Short Schwartz's Value Survey. *Journal of Personality Assessment, 85*, 170-178.
- Lönnqvist, J.-E., Leikas, S., Paunonen, S., Nissinen, V., & Verkasalo, M. (2006). Conformism moderates the relations between values, anticipated regret, and behavior. *Personality and Social Psychology Bulletin, 32*, 1469–1481.
- Lönnqvist, J.-E., Verkasalo, M., Helkama, K., Andreyeva, G. M., Bezmenova, I., Manganelli Rattazzi, A. M., Niit, T., & Stetsenko, A. (2007). Self-esteem and values. *European Journal of Social Psychology, 39*, 40-51.
- Lönnqvist, J.-E., Verkasalo, M., Wichardt, P., & Walkowitz, G. (2012). *Motivational Complexity Moderates the Relations between Personal Values and Behavior in Strategic Interactions*. Unpublished manuscript.
- Lönnqvist, J.-E., Verkasalo, M., & Walkowitz, G. (2011). It pays to pay – Big Five personality influences on co-operative behavior in an incentivized and hypothetical prisoner's dilemma game. *Personality and Individual Differences, 50*, 300-304.
- Lönnqvist, J.-E., Walkowitz, G., Verkasalo, M., & Wichardt, P. (2011). Situational power moderates the influence of Self-Transcendence vs. Self-Enhancement values on behavior in ultimatum bargaining. *Journal of Research in Personality, 45*, 336-339.

- Lönnqvist, J.-E., Walkowitz, G., Wichardt, P., Lindeman, M., & Verkasalo, M. (2009). The moderating effect of conformism values on the relations between otherpersonal values, social norms, moral obligation, and single altruistic behaviours. *British Journal of Social Psychology*, 48, 525–546.
- Maio, G. R., Olson, J. M., Bernard, M. M., & Luke, M. A. (2003). Ideologies, values, attitudes, and behavior. In J. DeLamater (Ed.), *Handbook of social psychology* (pp. 283-308). New York: Kluwer Academic.
- McCullough, B.D., McGeary, K. A., & Harrison, T. D. (2006) Lessons from The Journal of Money, Credit and Banking Archive. *The Journal of Money Credit and Banking*, 38, 1093-1107.
- Monin, B., & Merritt, A. (2011). Moral hypocrisy, moral inconsistency, and the struggle for moral integrity. In M. Mikulincer& P. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil.Herzliya Series on Personality and Social Psychology, Vol. 3*.American Psychological Association.
- Monin, B., Sawyer, P.J., & Marquez, M.J. (2008). The rejection of moralrebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, 95, 76–93.
- Myyry, L.,&Helkama, K. (2002). The role of value priorities and professional ethics training in moral sensitivity. *Journal of Moral Education*, 31, 35–50.
- Myyry L, Juujärvi S, Pesso K (2010). Empathy, perspective taking and personalvalues as predictors of moral schemas. *Journal of Moral Education*, 39, 213-233.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598 – 609.

- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality & Social Psychology, 81*, 524–539.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the big five inventory in English and German. *Journal of Research in Personality, 41*, 203–212.
- Rest, J. (1979). *Development in Judging Moral Issues*. University of Minnesota Press.
- Rocca, S., & Sagiv, L. (2010). Personal values and behavior: Taking the cultural context into account. *Social and Personality Psychology Compass, 4*, 30–41.
- Scheepers, D., Spears, R., Manstead, A. S. R., & Doosje, B. (2009). The influence of discrimination and fairness on collective self-esteem. *Personality and Social Psychology Bulletin, 35*, 505–515.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In Zanna, M. (Ed.), *Advances in Experimental Social Psychology, 25*, 1–65. New York: Academic Press.
- Schwartz, S. H. (2007). Universalism values and the inclusiveness of our moral universe. *Journal of Cross-Cultural Psychology, 38*, 711–728.
- Schwartz, S. H. (2009). Culture matters: National value cultures, sources and consequences. In C.-Y. Chiu, Y. Y. Hong, S. Shavitt, & R. S. Wyer, Jr. (Eds.), *Understanding culture: Theory, research and application*. New York: Psychology Press.

- Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology*, 32, 519–542.
- Schwartz, S. H., Verkasalo, M., Antonovsky, A., & Sagiv, L. (1997). Value priorities and social desirability: Much substance, some style. *British Journal of Social Psychology*, 36, 3–18.
- 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
- Silfver, M., Helkama, K., Lönnqvist, J.-E., Verkasalo, M. (2008) The relation between value priorities and proneness to guilt, shame and empathy. *Motivation and Emotion*, 32, 69–80.
- Valdesolo, P., & DeSteno, D. A. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology*, 44, 1334–1338.
- Verkasalo, M., Lönnqvist, J.-E., Lipsanen, J., & Helkama, K. (2009). European norms and equations for a two dimensional presentation of values as measured with Schwartz's 21-item portrait values questionnaire. *European Journal of Social Psychology*, 39, 780–792.
- Verplanken, B., & Holland, R. W. (2002). Motivated decision making: Effects of activation and self-centrality of values on choices and behavior. *Journal of Personality and Social Psychology*, 82, 434–447.
- Waples, E.P., Antes, A.L., Murphy, S.T., Connelly, S., & Mumford, M.D. (2009). A metaanalytic investigation of business ethics instruction. *Journal of Business Ethics*, 87, 133–151.

Wedekind, C., & Milinski, M. (2000). Cooperation through imagescoring in humans. *Science*, 288, 850–852.

Table 1. Dictator Behavior in Studies 1, 2, and 3.

	Outcome 8/2		Outcome 5/5	
Study 1	Without coin	With coin	Without coin	With coin
Watched	12 (38%)	16 (50%)	4 (13%)	0
Unwatched	18 (56%)	10 (31%)	4 (13%)	0
Collapsed	30 (47%)	26 (40%)	8 (13%)	0
Study 2				
Incentivized	26 (81%)		6 (19%)	
Hypothetical	15 (47%)		17 (53%)	
Study 3	22 (69%)	2 (6%)	6 (19%)	2 (6%)

Table 2. Morality Ratings of Dictator Behavior across different Groups (Study 1).

Choices	Passive person	Participants who chose:			Average	F value
		5/5 without coin	8/2 without coin	8/2 with coin		
5/5 without coin	7.33 _a	7.88 _a	7.07 _a	7.46 _a	7.33	0.72
5/5 with coin	5.78 _{a, b}	3.00 _{c, d}	4.67 _{a, d}	6.23 _b	5.44	6.23***
8/2 without coin	2.66 _a	0.75 _{b, c}	2.87 _a	1.77 _{a, c}	2.41	4.74**
8/2 with coin	4.60 _a	2.13 _{b, c}	4.03 _{a, c}	5.37 _a	4.46	4.67**
Coin effect	1.94 _a	1.38 _{a, b}	1.17 _a	3.58 _b	2.05	4.98**

Note. F-values are from one-way ANOVAs. Groups with different subscripts differ at $p < .05$ (Tukey HSD).

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

Figure 1.Schematic presentation of Studies 1, 2, and 3.

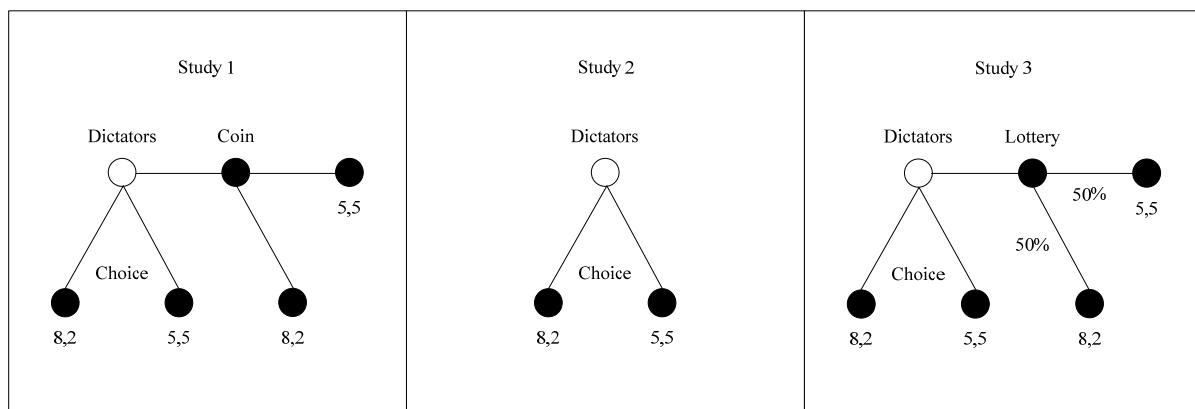


Figure 2. Fair vs. selfish choice regressed on low and high Conformity scores in the hypothetical and incentivized condition (Study 2).

Note: Low score = 1 SD below the mean; high score = 1 SD above the mean.

