Journal of Experimental Social Psychology 79 (2018) 71–77

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Daring dishonesty: On the role of sanctions for (un)ethical behavior

Isabel Thielmann*, Benjamin E. Hilbig

University of Koblenz-Landau, Germany

ARTICLE INFO

Keywords:
Cheating
Dishonesty
Sanctions
Costs
Punishment
Utility

ABSTRACT

According to seminal utility-based theories of norm-violating and unethical behavior, the decision to lie involves trading-off the potential benefits of dishonesty against the potential costs if caught. However, even in paradigms with zero risk of sanctions, individuals do not consistently cheat. Still more strikingly, most of the few findings available from studies implementing a sanctioning system run contrary to what would be expected based on utility-based accounts of dishonesty, showing increased cheating when there is a small risk of sanctions as compared to when the risk is zero. Given the striking nature of these findings and the general scarcity of corresponding evidence, we devised a targeted empirical test of the role of sanctions for dishonest behavior. Specifically, in two experiments, we manipulated the probability (i.e., 16.7% vs. 50%) that the factual truth of one's response in a cheating paradigm is checked. In Experiment 2, we further varied the severity of sanctions for being caught cheating. Results showed that the willingness to lie strongly decreased with increasing probability of being caught as well as with increasing severity of sanctions. As such, the results clearly support the theoretical notion that external costs influence unethical behavior, in the direction implied by utility-based models.

Lies corrupt society at all levels. At the individual level, people cheat on their partners, they evade taxes, and they commit insurance fraud. At the global level, in turn, large automobile groups fake emission protocols, politicians misappropriate public funds, and governments establish nationwide doping of athletes. Indeed, dishonest and unethical behaviors more generally can take many forms and typically cause considerable societal costs (Mazar & Ariely, 2006), thereby posing a severe threat to societal functioning. It is thus unsurprising that the study of dishonest behavior has attracted broad interest across scientific fields.

Early theories of norm-violating behavior rooted in economics (most prominently Becker, 1968) view dishonesty through the lens of expected utility models: Individuals are assumed to trade-off the potential utilities (benefits) of lying against its potential disutilities (costs) and their respective probabilities of occurrence. Whenever this cost-benefit analysis turns out positive, the normative response would be to lie. To illustrate this, consider a typical *cheating paradigm* as is commonly used to study dishonest behavior in experimental settings, the die-rolling paradigm (e.g., Fischbacher & Föllmi-Heusi, 2013; Shalvi, Handgraaf, & De Dreu, 2011). Participants are asked to roll a die in private and to report the outcome of their roll. They know that a certain number (originally a “5”) is associated with the highest payoff whereas the remaining numbers yield smaller (if any) payoffs. Importantly, participants also know that there is a zero risk of getting caught cheating because the experimenter does not check the truth status of participants' responses. Thus, lying is possible and leads to a certain gain without incurring any risk of sanctions in terms of material or social (i.e., reputation- or status-related) – and thus external – costs. According to a traditional utility-based account of dishonesty (Becker, 1968), individuals should therefore largely cheat by misreporting their actual outcome to obtain the maximum payoff.

In apparent contrast to this prediction, however, a consistent finding from the die-rolling and similar paradigms with zero risk of sanctions is that many individuals refrain from lying to the maximum extent but rather cheat a little, and some completely shy away from lying (e.g., Abeler, Nosenzo, & Raymond, 2016; Fischbacher & Föllmi-Heusi, 2013; Hilbig & Thielmann, 2017; Peer, Acquisti, & Shalvi, 2014). For example, in the die-rolling paradigm as previously sketched, a substantial proportion of participants report a 4 – rather than a 5 – which leads to the second highest payoff without requiring individuals to lie to the maximum extent (e.g., Fischbacher & Föllmi-Heusi, 2013; Grolleau, Kocher, & Sutan, 2016; Shalvi, Handgraaf, & De Dreu, 2011; see Hilbig & Hessler, 2013, for a similar approach). Such findings have been accounted for by assuming that lying may not only incur external costs, but also internal (i.e., psychological) ones: Whereas external costs subsume any type of negative consequences – both material (e.g., money) and immaterial (e.g., a loss in status) – that others might impose upon someone being caught cheating, internal costs are essentially

* Corresponding author at: Cognitive Psychology Lab, Department of Psychology, University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany.
E-mail address: thielmann@uni-landau.de (I. Thielmann).

https://doi.org/10.1016/j.jesp.2018.06.009
Received 15 March 2018; Received in revised form 19 June 2018; Accepted 28 June 2018

those imposed by the cheating individual herself. Specifically, prominent theories such as self-maintenance theory (Mazar, Amir, & Ariely, 2008) assume that any unethical behavior will, to some extent, pose a threat to one's moral self-image, undermining one's desire to perceive oneself as a "good person". Thus, importantly, internal costs occur independent of the risk of getting caught.

Further supporting the notion of internal costs, it has been shown that individuals treat gains obtained through cheating as though they were hard-earned (Thielmann & Hilbig, in press) and cheat more if this is more easily justifiable (e.g., Bassarak et al., 2017; Klein, Thielmann, Hilbig, & Zettler, 2017; Pittarello, Leib, Gordon-Hecker, & Shalvi, 2015; Shalvi, Dana, Handgraaf, & De Dreu, 2011; Shalvi, Gino, Barkan, & Ayal, 2015) – essentially allowing them to profit from cheating while maintaining their positive self-view. A strong implication from prior research investigating cheating in the absence of sanctions might thus be that individuals primarily trade-off the gains of lying against the *internal* (psychological) costs of self-image threat.

Importantly, however, these findings and corresponding explanations cannot be taken to imply that external costs are irrelevant for the decision to lie. Indeed, self-maintenance theory does not suggest that external costs are ineffective to prevent individuals from lying. Strikingly, though, almost all prior research has addressed the influence of internal costs for lying by relying on paradigms in which lying remains without any risk of sanctions (as in the die-rolling paradigm described above). As a consequence, most findings are actually mute on whether individuals might engage in a cost-benefit analysis involving external costs. In other words, although prior evidence shows that individuals mostly refrain from lying to the maximum extent even in the absence of external costs, it cannot be ruled out that individuals actually do trade-off external gains and costs once the latter are possible. Arguably, this is a vital question because many – if not most – real-life opportunities for dishonesty will be associated with some risk of getting caught and experiencing corresponding (monetary and/or social) sanctions such as fines, loss of reputation, or even jail sentences. Also, from a societal perspective, external costs (sanctions and their probability and severity) are the only realistic toehold for influencing the prevalence of unethical behavior.

Indeed, a few relevant empirical investigations are available on the role of sanctions for ethical behavior. Unfortunately, though, their findings are mixed, with the majority actually running contrary to what would be expected from a utility-based approach. That is, some studies found that lying *increased* when sanctions were implemented with comparably small probability. For example, Gamliel and Peer (2013) checked the truth status of participants' responses in a matrix task with 0%, 2.8%, or 100% probability. Counterintuitively, the (descriptively) highest level of cheating was apparent once there was a small (2.8%) risk of getting caught. Similarly, when individuals faced a 0%, 20%, or 100% probability that the truth status of their response would be checked, they were *more* likely to cheat when there was some (as compared to a zero) risk of getting caught and experiencing any type of sanctions (Békir, Harbi, Grolleau, Mzoughi, & Sutan, 2016). However, it should be noted that comparisons in this latter study were hampered by limited statistical power, as each experimental condition counted only around 25 participants. Nonetheless, these findings are overall in line with studies showing a positive link between cheating and sensation seeking (DeAndrea, Carpenter, Shulman, & Levine, 2009; McTernan, Love, & Rettinger, 2014), suggesting that lying may be associated with a certain thrill, which is arguably increased whenever sanctions are possible.

By contrast, however, there is also evidence indicating lower levels of cheating when participants were informed that an unknown proportion of responses in a coin-tossing task would be checked and that participants might lose their payment in case they cheated (Jones & Paulhus, 2017). Likewise, whenever participants know for sure that their responses will be checked, they usually refrain from lying (e.g., Gino, Ayal, & Ariely, 2013; Grolleau et al., 2016; Mazar et al., 2008).

Taken together, these prior studies provide some insights into the influence of sanctions on dishonest behavior. However, evidence in this regard is rather inconsistent and only rudimentary in general. That is, very few studies have implemented a sanctioning system in cheating paradigms at all, and those studies implementing sanctions have either only compared a certain (small) risk of sanctions against a zero and perfect risk (Békir et al., 2016; Gamliel & Peer, 2013) or an unknown risk of sanctions against a zero risk (Jones & Paulhus, 2017). Evidence on the effect of different (larger than 0 and smaller than 1) probabilities that one's response may be checked is thus missing entirely. In addition, it is still unclear to what extent the risk *and* severity of sanctions may uniquely affect the decision to lie. Most importantly, however, the majority of extant findings actually imply that an increase in external costs will *not* reduce dishonesty but may rather *increase* it. Given the extreme practical (policy) implications of such findings suggesting that "a fine is a price" (Gneezy & Rustichini, 2000) when it comes to unethical behavior – meaning that individuals may feel that unethical behavior is justified *because* they will potentially pay for it – additional tests are arguably vital. The present work aimed at addressing this issue by studying the effect of risk and severity of sanctions on cheating behavior in two experiments. Thereby, we provide a direct test of whether the probability and/or severity of external costs influence the prevalence of dishonest behavior.

In methodological terms, note that we report how we determined our sample sizes, all data exclusions (if any), all manipulations, and all measures in both experiments. The data was analyzed only after completing data collection corresponding to the a priori sample size calculations (see below). The materials as well as the data are available online at the Open Science Framework (OSF; https://osf.io/zkfhs/). We stress that no deception was used throughout the experiments, in line with common ethical guidelines.

1. Experiment 1

Experiment 1 was designed to investigate the effect of different risks of sanctions on cheating in a dice task. Specifically, participants could lie to obtain a (monetary) gain while either facing a 16.7% or a 50% probability that the truth status of their response would be checked. In case they were found to have cheated, participants lost their gain and part of their endowment obtained in another unrelated task. Thus, cheating was possible, but associated with considerable costs with small to moderate probability.

1.1. Method

1.1.1. Design, materials, and procedure

The experiment was conducted at the laboratory of a German university as part of a larger study, and it followed an unrelated decision-making task (i.e., a series of social dilemma games). Participants first received detailed (written and oral) instructions on the dice task (see details in the OSF supplement) in which participants could give dishonest responses to increase their monetary gain. The task consisted of two rounds with varying probabilities p_c that the truth status of a participant's response would be checked (within-subjects): In the *low risk condition*, responses were checked with 16.7% probability whereas in the *high risk condition* they were checked with 50% probability. In case a participant was found to have cheated in a given round, this incurred a net loss of −1€ (see details below). Fig. 1 provides an overview of the experimental task.

Participants were informed that they would play two rounds of a task in which they could gain up to 10€ (approx. US$ 12.90 at the time of data collection), but also lose up to 2€ from their payment obtained in the previous, unrelated task. Participants completed the dice task one by one in a separate room with the experimenter. First, they were asked to determine their *target number* by drawing a card from a deck containing six cards with numbers 1 to 6. Participants determined their