

EMPIRICAL ARTICLE

When and why does observability increase honesty? The role of gossip and reputational concern

Annika S. Nieper¹ , Bianca Beersma¹, Maria T. M. Dijkstra¹, and Gerben A. Van Kleef²

¹Vrije Universiteit Amsterdam, Amsterdam, Netherlands and ²Universiteit van Amsterdam, Amsterdam, Netherlands

Corresponding author: Bianca Beersma; Email: b.beersma@vu.nl

Received: 13 December 2022; **Revised:** 12 March 2024; **Accepted:** 15 March 2024

Keywords: gossip; reputation; observability; unethical behavior; cheating; honesty

Abstract

People frequently engage in dishonest behavior, which entails costs to society. A common advice to increase honesty is to enhance observability. However, previous research produced conflicting findings, making it unclear when and why observability increases honesty. Here we show that observability enhances honesty when observers can gossip to relevant others (i.e., to future interaction partners who can influence the gossip target's outcome), because it increases reputational concern. In 2 incentivized and pre-registered studies, participants privately rolled a die 30 times and were informed that reporting higher numbers would lead to higher outcomes (total $N = 1608$; 28650 observations). We manipulated observability and gossip. Both studies revealed that gossip to relevant others decreased dishonest reporting, whereas mere observation did not. Importantly, reputational concern partly mediated the impact of gossip on dishonesty. Moreover, gossip influenced recipients' trust in gossip targets, with messages denoting dishonesty swaying trust more than messages denoting honesty. Our findings demonstrate when and why observability promotes honesty.

Everyday life presents ample opportunities for obtaining personal gains by behaving dishonestly, for instance by incorrectly filling out one's tax reports or declaring higher personal expenses after a business trip. Such dishonest acts entail substantial societal costs (Slemrod, 2016). Leaders and policy makers therefore implement policies to increase honest behavior. Many such policies involve making behavior observable to others. For example, in 2001, Norway implemented a system of public disclosure of tax and income information to increase tax compliance. Since then, neighbors and acquaintances can view other people's income and tax reports online (Bø et al., 2015). Similarly, companies often implement a 4-eyes principle, which means another person can observe how one has behaved, to increase honesty (Bodenschatz and Irlenbusch, 2019).

Despite the popularity of these approaches, previous studies on the effects of observability on honesty have yielded inconsistent results. On the one hand, there are several studies that show that observability increases honesty (Abeler et al., 2019; Bø et al., 2015; Bodenschatz and Irlenbusch, 2019; Houser et al., 2016; Köbis et al., 2019; Lilleholt et al., 2020). On the other hand, there are also several studies that did not find evidence that manipulating observability influences honesty (Abeler et al., 2019; Bodenschatz and Irlenbusch, 2019; Fischbacher and Föllmi-Heusi, 2013; Pfaffheicher et al., 2019). In addition, Weisel and Shalvi (2015) found that in a task in which people could be dishonest together, they behaved more dishonestly in comparison to an individual condition (in which they were

not observed). Likewise, Bäker and Metchel (2019) found more cheating in a setting in which peers were present in comparison to a setting where no peers were present. Thus, the latter 2 studies indicate that observability could also increase dishonesty.

The above overview shows that the relationship between observability and dishonest behavior is inconsistent; several studies found that observability decreases dishonesty, whereas others found no relationship between observability and dishonesty, and 2 studies even found that observability increases dishonesty. These disparate findings present a puzzle that is yet unresolved: When and why does observability increase honesty? We propose the effects of observability on honesty depend on whether being observed increases reputational concern, which we study by examining opportunities for gossip, because gossip increases reputational concern.

1. Gossip and reputational concern

Reputational concern—concern about the estimation others hold about one’s qualities or characteristics (Caldwell, 1986)—is an important driver of human behavior (Abeler et al., 2019; Engelmann and Rapp, 2018). We propose that observability increases honesty to the degree that people are concerned about their reputation. People are not equally concerned about their reputation in all situations. Reputational concern is particularly high when people are aware that interaction partners can transmit information about them to their future interaction partner(s); Beersma and Van Kleef, 2011). Such information is commonly transmitted via gossip, defined as the exchange of information about absent others (Dores Cruz et al., 2021; Fonseca and Peters, 2018). Gossip consists of a gossip triad, comprised of at least 3 people: those who send it (gossip sender), those who receive it (gossip receiver) and those who are gossiped about (gossip target; see Dores Cruz et al. (2021) for an integrative review of gossip definitions and a detailed discussion about what constitutes of gossip or not). People preferentially gossip to others within their social networks, that is, to potential future interaction partners of the gossip target (Robbins and Karan, 2020) who can benefit or harm the target’s outcomes (Dores Cruz et al., 2020).

Accordingly, previous research has demonstrated that gossip can increase reputational concern and thereby cooperation. People are more cooperative when others can gossip about them to relevant others (Peters and Fonseca, 2020; Wu et al., 2015), because the possibility of gossip to relevant others increases reputational concern (Beersma and Van Kleef, 2011; Wu et al., 2015). Although this prior work does not speak to the possible effects of gossip on honesty, it does point to the importance of reputational concern.

The prior findings about the relationship between reputational concerns and cooperation, however, do not allow drawing conclusions about how reputational concern affects honesty. Cooperation differs significantly from honesty on several dimensions. First, when people make the decision to cooperate or not, those decisions always impact another person’s outcome, because cooperation by definition involves interdependent others (West et al., 2006). In contrast, when deciding to be honest or not, those decisions can but do not have to impact another person’s outcome (Abeler et al., 2019; Leib et al., 2021; Schild et al., 2020). Second, people need to actively engage in dishonesty, while in order to defect (i.e., not cooperate), people are simply withholding cooperation and do not actively have to engage in a different type of behavior. Third, dishonesty unambiguously falls into the realm of moral decisions, while cooperative decisions are not always considered as moral decisions. Lastly, the relationship between reputational concern triggered by gossip and dishonest behavior may be a less obvious one than the relationship between gossip and cooperation, as gossip tends to occur in situations in which people are affected by a norm-violator’s outcomes (Beersma and Van Kleef, 2011). As stated, this is the case for non-cooperation, but it need not be the case for dishonesty. Therefore, whereas people may have evolved or learned (Wu et al., 2016) to avoid gossip about them by cooperating, this pattern may not be as well established for engaging in honest behavior. In conclusion, cooperation is related to, but different from (dis)honesty, and whereas the previously established link between reputational concerns and cooperation helps to make predictions about the impact of reputational concerns on (dis)honesty, predictions related to this need to be tested empirically.

Here, we propose that reputational concern can reconcile the previous conflicting findings regarding the effects of observability on dishonesty, and test this idea by varying reputational concerns by manipulating gossip, because previous research has shown that gossip impacts people's reputational concerns. We predict that people behave more honestly when they are observed by another person who has the possibility to gossip about them to relevant others than when they are not observed or the observer has no possibility to gossip, because the combination of observation and gossip to relevant others increases reputational concern.

2. Gossip and trust

Although our primary focus is on how gossip directly encourages *momentary* honesty, gossip may also indirectly regulate honesty by shaping trust—a foundation of cooperative social relationships (Balliet and Van Lange, 2013). When observers can gossip, they may send gossip in which they describe a target as either honesty or dishonesty to recipients. Gossip recipients may in turn condition their trust on the nature of the information they receive, such that they exhibit greater trust in those about whom they received information denoting honesty rather than dishonesty (Pagliaro et al., 2013). If this occurs, an overall effect of gossip on trust can, but does not need to, occur.

3. Overview of the current research

We tested our core idea—that observability enhances honesty when observers can gossip to relevant others—in 2 pre-registered behavioral experiments. In Study 1, we tested the predictions that people behave more honestly when they know they are being observed and observers can gossip about them to relevant others, compared to when they are not observed (*Hypothesis 1a*) and when they are observed by someone who cannot gossip (*Hypothesis 1b*). We expected the influence of gossip on honesty to be mediated by reputational concern (*Hypothesis 2*). Besides testing these pre-registered hypotheses, we explored the impact of gossip on subsequent trust by examining the overall impact of gossip on trust and examining whether gossip recipients condition their trust on the nature of the gossip they received. In Study 2, we aimed to replicate the findings of Study 1 and included a gossip condition in which the gossip was sent to irrelevant (instead of relevant) others to illuminate a boundary condition of the effect of gossip on honesty. In both studies, we examined real interactions, did not use any deception and incentivized participants for their decisions.

4. Study 1

4.1. Methods

4.1.1. Experimental design

We examined (dis)honest behavior using a validated die-rolling task (Fischbacher and Föllmi-Heusi, 2013; Shalvi et al., 2011). In this task, participants were asked to privately roll a die 30 times and report the outcomes, with higher reported values yielding higher payments. Specifically, die rollers earned experimental credit units (ECU) for each pip they reported (i.e., 1 ECU for reporting a 1; 6 ECU for reporting a 6). One ECU was worth 1 dollar cent, and all earned ECUs were converted into a bonus payment. Thus, there was a financial incentive to report higher outcomes than actually obtained. Participants were asked to either use a fair physical die or use a website such as '<https://freeonlinedice.com/>', which rolls a die at random. As such, the experimenters did not learn about the actual die rolls (so-called 'ground truth'), and to measure dishonesty, reported outcomes were compared against a chance-level average. Based on honest reporting (i.e., chance-level), one would expect an average of around 3.5 across the 30 die rolls; significantly higher averages indicate that people reported higher numbers than expected.

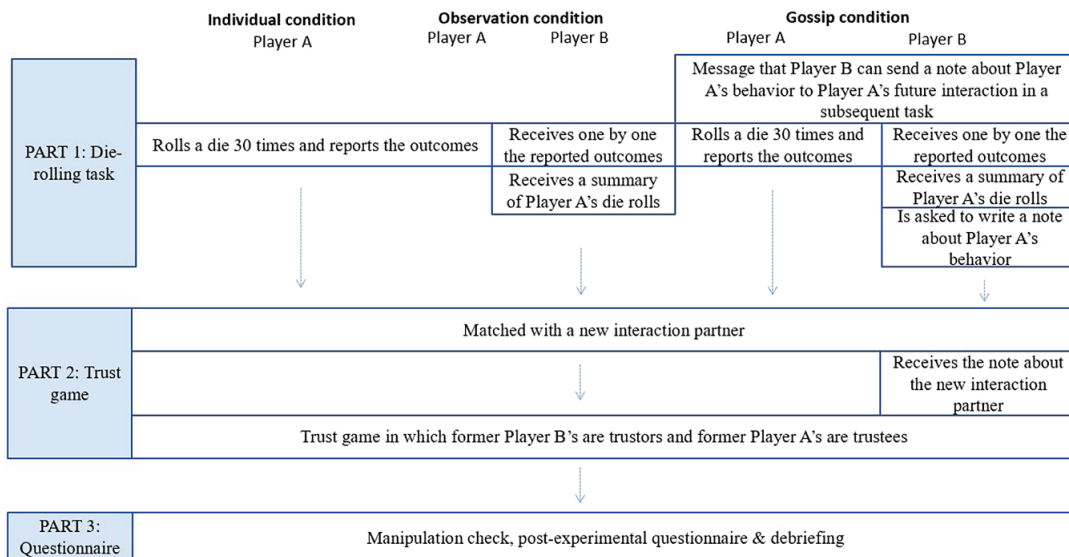


Figure 1. Overview of the conditions, roles, and parts of the experiment.

We compared participants’ reports across 3 conditions in a between-subjects design: (1) an individual condition, (2) an observation condition, and (3) a gossip condition. In the observation and gossip condition, participants were randomly assigned to either the role of Player A—the die roller—or Player B—the observer who received the die roll outcomes Player A reported. In the gossip condition, Player B was asked to write a note about the die roller’s behavior after receiving all reported die rolls using a free-form text field, that is, participants could write their own messages. As the note senders (just like the experimenters) did not have access to actual (‘ground truth’) information on whether participants lied or reported the die rolls truthfully, they could only make inferences based on the die roller’s behavior. This note was sent to the die roller’s future interaction partner in the second part of the study, a trust game (Berg et al., 1995). The first author and a research assistant independently coded the gossip statements for references to (dis)honesty.¹

The experiment consisted of 3 parts: the die-rolling task (Part 1), a trust game (Part 2), and a questionnaire (Part 3). Participants received instructions for each part of the experiment at the beginning of the respective part. Thus, although during Part 1 participants knew they would interact with someone else in a second task, they were not aware of the nature of this task, and therefore die rollers in the gossip condition could not take the nature of the task in which they would interact with their future partner into account. At the end, participants completed a manipulation check and a post-experimental questionnaire. Figure 1 provides a schematic illustration of the different conditions, roles, and parts of the experiment.

4.1.2. Procedure

After receiving approval from the Ethics Review Board, we conducted the experiment using the Software Platform for Human Interaction Experiments (SoPHIE; Hendriks, 2012), which allows online real-time interactions. The experiment took approximately 25 minutes.

¹Please note that in all conditions, participants’ reports could be observed by the experimenter. Thus, the individual condition could be considered an “observability by experimenter only” condition. However, we did not mention this experimenter observability explicitly to the participants in none of the conditions. This stands in contrast to the observability condition, where we explicitly instructed participants that their reported outcomes would be observed by another participant. We therefore do not emphasize the observability by the experimenter in our condition label.

4.1.3. Trust game

After Part 1 was completed, participants were matched with a new partner for Part 2, which consisted of a trust game (Berg et al., 1995). The trust game is a decision-making task in which one person, the trustor, can decide to send an amount of money to another person, the trustee. The transferred money is tripled and the trustee can decide to send any amount between zero and the tripled amount back to the trustor. In this task, trust is measured as the amount the trustor sends and trustworthiness is measured as the percentage the trustee sends back to the trustor (Johnson and Mislin, 2011). If a trustor did not send any money, we could not measure trustworthiness for that specific pair. Therefore, only participants for whom a trustworthiness score was available were included in the analyses.

In the individual condition, the participant was randomly assigned the role of the trustor or trustee, and in the gossip and observation conditions, the former Players B (observers) were the trustors, while the former Players A (die rollers) were the trustees. In the gossip condition, a Player B received the gossip about their new interaction partner from the Player B who observed this interaction partner during the die-rolling task. The trustor received 10 ECU and could decide to transfer any value between 0 ECU and 10 ECU to their newly matched interaction partner. Each ECU earned was worth \$0.01 and converted to a bonus payment at the end of the experiment.

4.1.4. Post experimental questionnaire

We measured reputational concern to test Hypothesis 3. Participants indicated their agreement with 9 questions on Likert-scales anchored (1) *strongly disagree* to (7) *strongly agree* (adapted from Beersma and Van Kleef, 2011). Example items include: ‘I thought about how a direct observer would think about me’ and ‘It was important to me that a direct observer would perceive me in a positive way’ (Cronbach’s alpha = .95). Higher scores on this scale reflect greater reputational concern.

4.1.5. Gossip statements

The first author and a research assistant independently coded the gossip statements for references to (dis)honesty. Statements were coded as portraying the participant as dishonest, honest, or unclear with regards to honesty. For example, we coded the statement ‘Player A seemed to be getting higher numbers on their die a lot more than what they should have’ as indicating that Player A was dishonest, the statement ‘Player A appears to be honest and reporting his die rolls accurately’ as indicating that Player A was honest, and the statement ‘They did a great job’ as unclear. There was substantial agreement between the 2 coders (Cohen’s kappa = .76; McHugh, 2012) and any disagreement was resolved through discussion. The results regarding trust and trustworthiness lead to similar results and identical conclusions when we analyze the initial coding of Coder 1 or Coder 2, rather than the coding in which inter-coder disagreements had been resolved through discussion (see the RMarkdown file on OSF for the results of these analyses and see [Supplementary Material](#) for an overview of all the gossip statements and how they were coded).

4.1.6. Participants

We recruited participants from the online platform MTurk, which yields more representative data compared to recruiting participants from university campuses (Casler et al., 2013). To ensure high-quality data, we employed the following procedures. For Study 1, for which we recruited participants on MTurk, in order to avoid pool contamination, we included a Captcha to prevent participation by robots. We further applied the following restrictions: Participants had to be in the U.S., have an approval rate greater than 99%, and a minimum of 1000 approved HITs. Participants received a fixed pay-off of \$3 for their participation, which reflects a common minimum hourly wage (~\$7.25/hour). In addition, die rollers earned an average bonus of \$1.18 based on their reported outcomes. Based on an a-priori power analysis, we sampled at least 160 participants per condition and role (800 participants) to ensure sufficient power (.80) to detect a small to medium effect size of 0.3 (Cohen’s *d*). In total, 907 participants fully completed the study (463 males, age $M = 37.51$ years, $SD = 10.97$ years, 14670

observations).² 227 participants were in the individual condition, 348 in the observability condition (174 die rollers), and 332 in the gossip condition (166 die rollers). No participants who fully completed the study were excluded.³

4.1.7. Analyses

Hypotheses 1 to 3 and their confirmatory tests were pre-registered at the Open Science Framework, and the data, code, and materials are available online (https://osf.io/verzn/?view_only=4b0bc2be8f964872af679c1a1f178324). We tested the directional pre-registered hypotheses with one-sided tests (Lakens, 2016), and all other effects with 2-sided tests.

4.2. Results

4.2.1. Confirmatory analyses

4.2.1.1. (Dis-)honesty levels

Because we have multiple non-independent observations (reported outcomes of 30 die rolls) per participant, we used multilevel regression to test our main hypotheses (Bates et al., 2019). To account for the nested structure of our data (observations nested within participants), we specified one intercept for each participant (with multiple observations for all 30 die rolls). Based on honest reporting (i.e., chance-level), one would expect an average of around 3.5 across the 30 die rolls; significantly higher averages indicate that people reported higher numbers than expected.

In the gossip condition participants' scores were closest to what would be expected based on honest reporting (Figure 2). Specifically, supporting H1, participants were more honest in the gossip condition ($M = 3.81$, 95% CI [3.73, 3.90]) than in the individual condition ($M = 3.93$, 95% CI [3.85, 4.02]), $b = -0.12$, 95% CI [-0.22, -0.02], $p = .026$, one-sided. Furthermore, supporting H2, participants were more honest in the gossip condition than in the observation condition ($M = 4.02$, 95% CI [3.93, 4.11]), $b = -0.21$, 95% CI [-0.31, -0.10], $p < .001$, one-sided. The individual and observation conditions did not differ significantly from each other, $b = 0.04$, 95% CI [-0.02, 0.11], $p = .173$.

4.2.1.2. Reputational concern

Participants reported higher reputational concern in the gossip condition ($M = 3.83$, 95% CI [3.57, 4.08]) than in the observation condition ($M = 3.27$, 95% CI [3.01, 3.54]), $t(338) = 2.97$, $d = 0.32$, $p = .002$, one-sided), and marginally higher reputational concern in the gossip condition than in the individual condition ($M = 3.53$, 95% CI [3.27, 3.78]), $t(382) = 1.62$, $d = 0.16$, $p = .050$, one-sided). The observation and individual conditions did not differ from one another, $t(389) = -1.37$, $p = .173$, $d = 0.14$ (Figure 3). Furthermore, reputational concern predicted honesty levels across conditions, $b = -0.08$, 95% CI [-0.11, -0.04], $p < .001$. We used the mediation package in R (Tingley et al., 2014) to test for mediation via reputational concern. Supporting H3, reputational concern mediated the contrast between the individual and the gossip condition. Estimates based on 10,000 bootstrap samples revealed that the indirect effect was significant, $b = 0.02$, 95% CI [0.01, 0.03], $p = .002$. The direct effect of condition on (dis-)honesty became non-significant, $b = -0.14$, 95% CI [-0.29, 0.00], $p = .051$, indicating full mediation. Furthermore, reputational concern partially mediated the contrast between the gossip and observation condition. The indirect effect was again significant, $b = 0.04$, 95% CI [0.01, 0.08], $p = .003$, but the direct effect of condition on (dis-)honesty remained significant, $b = 0.23$, 95% CI [0.07, 0.38], $p = .004$, indicating partial mediation.

²In Study 1, in two sessions, we had synchronization issues due to which participants could not be matched with another person to play the trust game with. These sessions were not included in the count of participants who fully completed the study. Furthermore, some participants only partially completed the experiment because they dropped out of the study while waiting for their partner's responses. In case these participants provided data pertaining to the relevant hypotheses, we included their responses in the tests of H1–H3.

³Below, we report the results of analyses on the full sample. As pre-registered, we also analyzed the data of only those participants who filled in the manipulation checks correctly. These analyses are reported in the [Supplementary Material](#). Importantly, both sets of analyses lead to the same results and conclusions.

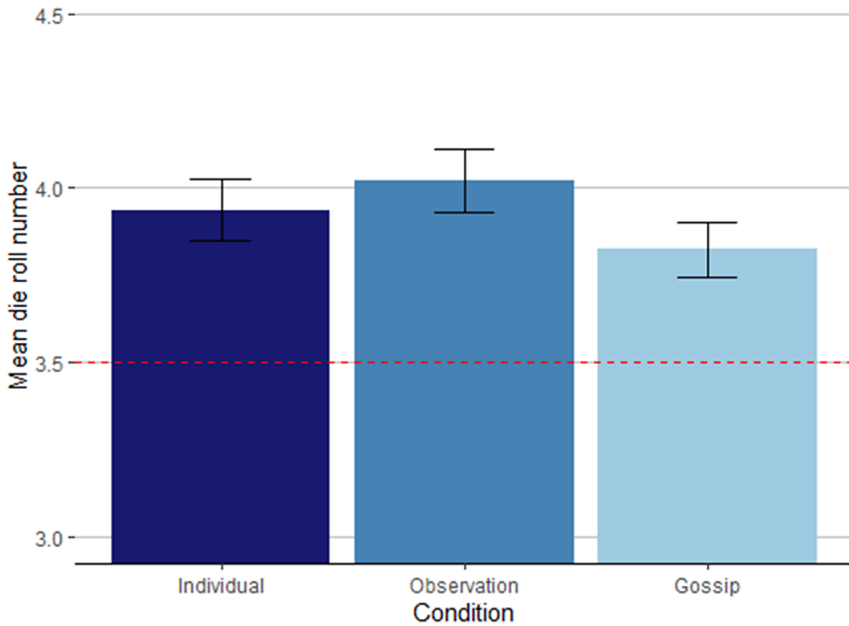


Figure 2. Average reported die roll outcomes per condition. The red dotted line indicates the expected mean die roll outcome based on chance. Errors bars show 95% CIs.

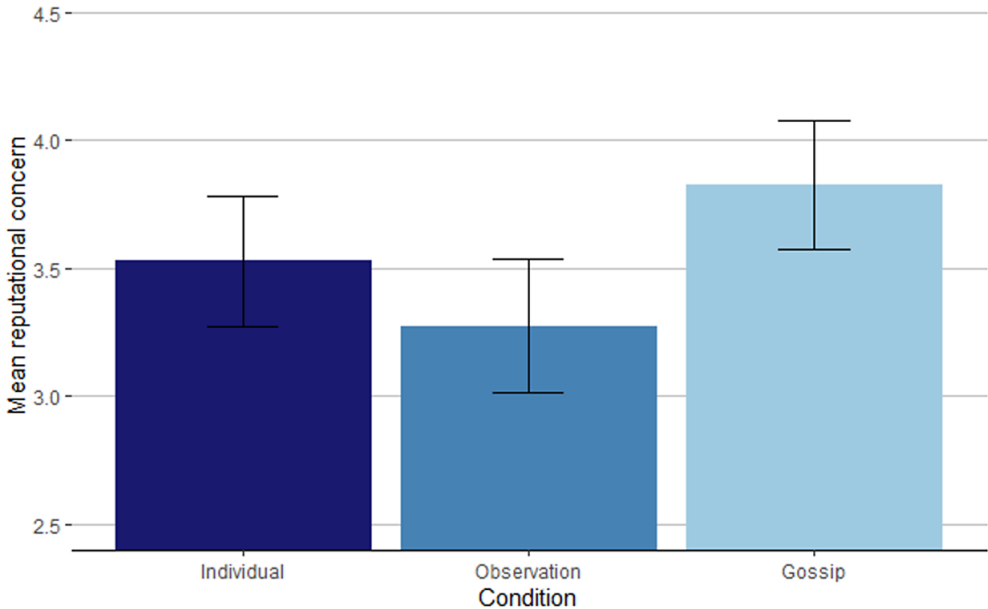


Figure 3. Reputational concern per condition. Errors bars show 95% CIs.

4.2.2. Exploratory analyses

4.2.2.1. Do people still cheat when they can be gossiped about?

The analyses reported above test whether there are significant differences between the gossip and the observability condition, and the gossip and the individual condition. That is, they test for a main effect of condition. We furthermore test exploratively whether the die roll distributions in the 3 experimental

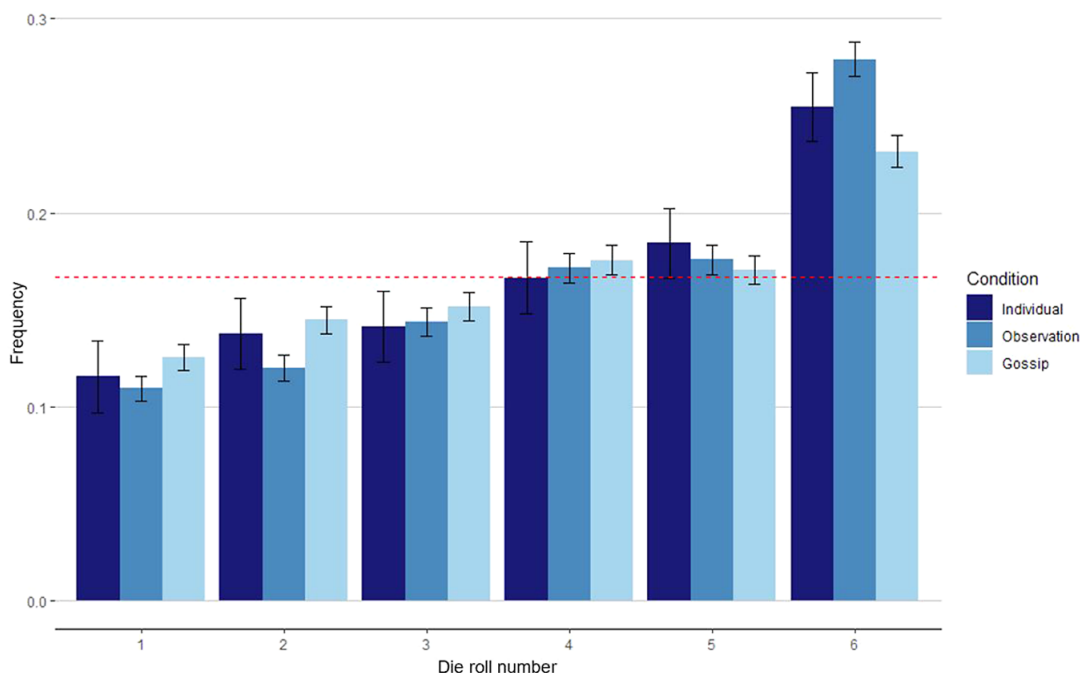


Figure 4. Distribution of reported die roll outcomes by condition. The dotted red line indicates the frequency based on chance.

conditions differ from a uniform distribution, which would be expected based on chance when people report honestly. These findings give insights into whether on average people still overreported in these conditions, or not. The reported die roll outcomes in all conditions were significantly different from a uniform distribution (i.e., chance level) according to a Kolmogorov–Smirnov test, $D_s > .23$, $p_s < .001$ (Figure 4). This indicates that, on average, people in all conditions over-reported their die roll outcomes (albeit significantly less so in the gossip condition, as predicted).

4.2.2.2. Gossip

In the majority ($N = 114$; 58%) of all gossip statements ($N = 198$), die rollers were portrayed as honest. In 34 gossip messages, die rollers were portrayed as dishonest, and 50 statements were unclear. There was a strong, significant correlation between dishonesty levels (the die roll average) and gossip content (honest vs. dishonest), $r = .77$, 95% CI [.70, .83], $p < .001$. The higher the die roll averages, the more likely the person was portrayed as dishonest.

4.2.2.3. Trust

We tested whether gossip recipients conditioned their trust on the content of the gossip received (dishonest vs. honest statements). Indeed, people trusted others more when they had received gossip that their interaction partner was honest ($M = 6.59$, 95% CI [5.90, 7.29]) rather than dishonest ($M = 3.42$, 95% CI [2.01, 4.78], $b = -3.17$, 95% CI [-4.66, -1.69], $p < .001$) in the previous task. Notably, this effect was driven more by Player B's negative reaction to the information that Player A was dishonest than by Player B's positive reaction to the information that Player A was honest. That is, Player B trusted Player A similarly when receiving no message ($M = 5.99$, 95% CI [5.53, 6.44]) and when receiving a message suggesting the other was honest ($M = 6.59$, 95% CI [5.90, 7.29], $b = 0.60$, 95% CI [-0.26, 1.47], $p = .172$), but Player B trusted Player A significantly less when receiving a message suggesting the other was dishonest ($M = 3.42$, 95% CI [2.01, 4.78]) than when receiving no message at all, $b = -2.57$, 95% CI [-4.01, -1.13], $p < .001$. Thus, that interaction partners are honest and

can therefore be trusted seems to be the default presumption and this changes when gossip informs recipients otherwise.

4.2.2.4. Trustworthiness

Trustworthiness differed depending on the type of gossip received. Participants who were depicted as honest returned a higher percentage of the ECUs they received ($M = 0.41$, 95% CI [0.37, 0.45]) compared to participants who were described as dishonest ($M = 0.23$, 95% CI [0.151 0.36], $b = -0.18$, 95% CI [-0.28, -0.07], $p = .001$), indicating that the former were indeed more trustworthy than the latter. Lastly, we explored whether the die roll average predicted trustworthiness. We found a significant correlation between participants' die roll averages and their trustworthiness, $r = -.37$, 95% CI [-0.46, -0.28], $p < .001$ —the higher the die roll outcomes, the less trustworthy was the person.

5. Study 2

The objective of Study 2 was twofold. First, we wanted to replicate Study 1's hypothesized results (H1a, H1b, H2) as well as the exploratory findings that (1) trustors trust trustees less after receiving a message denoting that the trustee was dishonest rather than honest, (2) trustors trust trustees less after receiving a message denoting that the trustee was dishonest in comparison to receiving no message at all, and (3) trustees who are depicted as honest in gossip messages behave more trustworthily than trustees who are depicted as dishonest (pre-registration: https://osf.io/verzn/?view_only=4b0bc2be8f964872af679c1a1f178324).

Second, we wanted to shed additional light on the mechanism underlying the effect of gossip on honesty by testing a potential boundary condition: expected indirect benefits (Wu et al., 2015). We argued that observability increases honesty when observers can gossip about a target to relevant others. Others are arguably more 'relevant' when they can influence one's future outcomes than when they cannot. To test this boundary condition, we included a condition in which people could not earn money in the trust game, and gossip recipients were therefore not relevant to gossip targets' future outcomes.

5.1. Methods

5.1.1. Participants

We recruited participants from the online platform Prolific with the requirement that participants have English proficiency. Participants received £3.50 for their participation (~£8.40/hour). In addition, die rollers earned an average bonus of £1.09 based on their reported outcomes. Based on an a-priori power analysis, we sampled 100 participants per condition and role to ensure sufficient power ($> .80$) to detect an effect size of 0.12 (the regression coefficient from Study 1 of the effect of gossip on honesty). In total, 894 participants fully completed the study (503 males, age $M = 27.98$ years, $SD = 10.43$ years, 13980 observations)¹. In total, 118 participants were in the individual condition, 305 in the observability condition (152 die rollers), 232 in the gossip to relevant others (116 die rollers), and 240 in the gossip to irrelevant others condition (120 die rollers). No participants who fully excluded the study were excluded.

5.1.2. Deviations from Study 1

This experiment was identical to Study 1 except for 2 changes. First, we included an additional condition in which the trust game was not incentivized. In this condition, participants did not earn monetary benefits in the trust game, so that they could not expect indirect benefits from behaving honestly in the gossip condition. This led to 4 conditions in total: (1) an individual condition, (2) an observation condition, (3) a gossip to relevant others condition, and (4) a gossip to irrelevant others condition. The difference between the latter 2 conditions was that in the gossip to irrelevant others, participants were informed that they could not earn a bonus in the second part of the experiment (they

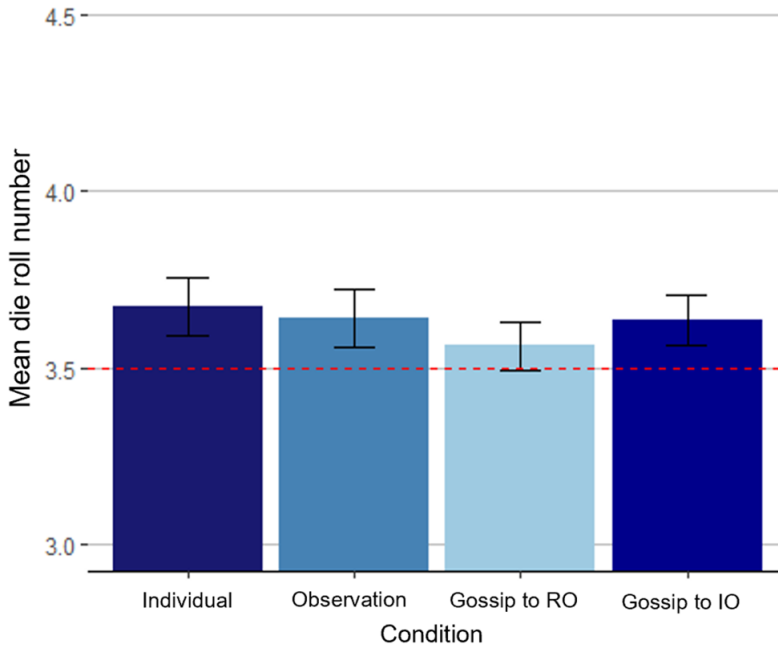


Figure 5. Average reported die roll outcomes per condition. The red dotted line indicates the expected mean die roll outcome based on chance. ‘Gossip to RO’ refers to the gossip to relevant others condition and ‘Gossip to IO’ refers to the gossip to irrelevant others condition. Error bars show 95% CIs.

could only earn money in the die-rolling task, not the second task). In the gossip to relevant others condition, just as in Study 1, participants could earn money in the die rolling task and the second task.

Second, we modified the procedure for gossip sending. In Study 1, participants could gossip via free-form text messages, which we coded for (dis)honesty by 2 trainers. In Study 2, to alleviate the need to have the data coded, we opted for having participants send a rating rather than a free-form message. Participants selected standardized gossip statements to send, and could choose from that ‘Player A seems to have (1) reported truthfully all the time, (2) reported truthfully most of the time, (3) reported truthfully half of the time and untruthfully half of the time, (4) reported untruthfully most of the time’, or ‘(5) reported untruthfully all the time’. We considered statements 1 and 2 as indicating honesty, statement 3 as mixed, and statements 4 and 5 as indicating dishonesty.

5.2. Confirmatory results

5.2.1. (Dis-)honesty levels

As shown in [Figure 5](#), in the gossip to relevant others (RO) condition, participants’ scores were closest to what would be expected based on honest reporting. Specifically, supporting H1a, participants were more honest in the gossip to RO condition ($M = 3.56$, 95% CI [3.50, 3.63]) than in the individual condition ($M = 3.67$, 95% CI [3.59, 3.76]), $b = -0.11$, 95% CI [-0.01, -0.22], $p = .02$, one-sided. Furthermore, partially supporting H1b, participants were somewhat more honest in the gossip to RO condition than in the observation condition ($M = 3.64$, 95% CI [3.56, 3.72]), $b = -0.08$, 95% CI [-0.18, 0.02], $p = .07$, one-sided. The individual and observation conditions did not differ significantly from each other, $b = -0.02$, 95% CI [-0.08, 0.04], $p = .565$, nor did any of the conditions from the gossip to irrelevant others (IO) condition ($M = 3.63$, 95% CI [3.57, 3.70]), $ps > .131$ ([Figure 5](#)).

5.2.2. Reputational concern

Participants reported higher reputational concern in the gossip to RO condition ($M = 4.23$, 95% CI [3.96, 4.51]) than in the observation condition ($M = 3.87$, 95% CI [3.62, 4.12]), $t(231) = 1.93$, $p = .028$,

one-sided), and in the individual condition ($M = 3.86$, 95% CI [3.60, 4.13], $t(239) = 1.90$, $p = .030$, one-sided). The observation and individual conditions did not differ from one another, $t(226) = -0.03$, $p = .973$. We used the mediation package in R (Tingley et al., 2014) to test for mediation via reputational concern. Not supporting H3, reputational concern did not mediate the effect of condition for the contrast between the individual and the gossip to RO condition nor between the observation and gossip to RO condition. Estimates based on 10,000 bootstrap samples revealed that the indirect effect was non-significant, $b = 0.00$, 95% CI [-0.02, 0.01], $p = .530$ for the contrast of gossip to RO versus observation, as well as for the contrast of gossip to RO versus individual, $b = 0.02$, 95% CI [0.00, 0.04], $p = .115$. In the gossip to irrelevant others condition, participants did not report higher reputational concern ($M = 4.12$, 95% CI [3.88, 4.37]) than in the observation condition, $t(222) = -1.42$, $p = .078$, one-sided, and did not report higher reputational concern than in the individual condition, $t(229) = -1.40$, $p = .081$. There was no indirect effect of reputational concerns for either contrast $ps > .31$.

5.2.3. Gossip

In the majority ($N = 189$; 79%) of all gossip statements ($N = 238$), die rollers were portrayed as honest. In 25 gossip messages, die rollers were portrayed as dishonest, and 24 statements were mixed. There was a significant correlation between dishonesty levels (the die roll average) and gossip message sent, $r = .24$, 95% CI [.12, .36], $p < .001$. The higher the die roll averages, the more likely the person was portrayed as dishonest.

5.2.4. Trust

We tested whether gossip recipients conditioned their trust on the content of the gossip received, and indeed the type of message received predicted recipients' trust toward the gossip target, $b = 0.59$, 95% CI [0.18, 0.99], $p = .005$. The more positive the gossip message, the more was the gossip target trusted. Notably, as in Study 1, this effect was driven more by Player B's negative reaction to the information that Player A was dishonest than by Player B's positive reaction to the information that Player A was honest. That is, Player B trusted Player A not significantly differently when receiving no message and when receiving a message suggesting Player A was honest, $b = 0.17$, 95% CI [-0.39, 0.73], $p = .552$, but Player B trusted Player A significantly less when receiving a message suggesting Player A was dishonest than when receiving no message at all, $b = -1.41$, 95% CI [-2.72, -0.10], $p = .017$, one-sided. Thus, as in Study 1, that interaction partners are honest and can therefore be trusted seems to be the default presumption and this changes when gossip informs recipients otherwise.

5.2.5. Trustworthiness

Similar to Study 1, trustworthiness differed depending on the type of gossip received about the target. Participants who were depicted as more honest returned a higher percentage of the ECUs they received compared to participants who were described as dishonest, $b = 1.50$, 95% CI [0.61, 2.38], $p < .001$, one-sided, indicating that the former were indeed more trustworthy than the latter.

5.3. Exploratory results

5.3.1. Do people still cheat?

The reported die roll outcomes in all conditions were significantly different from a uniform distribution (i.e., chance level) according to a Kolmogorov–Smirnov test, $Ds > .18$, $ps < .001$. This indicates that, on average, people in all conditions over-reported their die roll outcomes (i.e., reported higher outcomes than they actually rolled), albeit significantly less so in the gossip to relevant others condition (Figure 6).

6. Pooled data analysis

Both studies have 3 common conditions: (1) individual, (2) observation, and (3) gossip to RO condition. This makes it suitable to combine both datasets and summarize the effect across both datasets. Across both datasets, participants were more honest in the gossip to RO condition ($M = 3.73$ 95%CI [3.67,

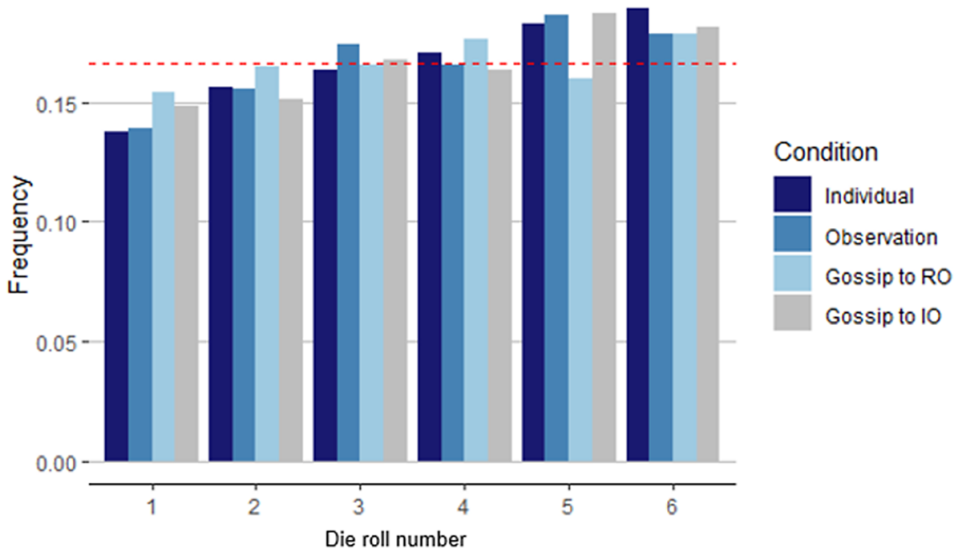


Figure 6. Distribution of reported die roll outcomes by condition. The dotted red line indicates the frequency of each outcome based on chance. Gossip to RO refers to the ‘gossip to relevant others’ condition and Gossip to IO refers to the ‘gossip to irrelevant others’ condition.

3.79]), in comparison to the individual condition ($M = 3.86$, 95%CI [3.79, 3.92]), $b = -0.126$, 95%CI [-0.216, -0.035], $p = .007$, and the observation condition ($M = 3.91$, 95% CI [3.84, 3.98]), $b = -0.179$, 95% CI [-0.272, 0.085], $p < .001$. Reputational concern partly mediated the impact of gossip on honesty for the contrast between the observation and gossip to RO condition. Estimates based on 5000 bootstrap samples revealed that the indirect effect was significant, $b = 0.03$, 95% CI [0.01, 0.05], $p < .001$. The direct effect of condition on (dis-)honesty was still significant, $b = 0.18$, 95% CI [0.07, 0.28], $p = .002$, indicating partial mediation. Similarly, reputational concern partially mediated the effect of condition for the contrast between the gossip to RO and individual condition. The indirect effect was significant, $b = -0.01$, 95% CI [-0.03, 0.00], $p = .013$, but the direct effect of condition on (dis-)honesty remained significant, $b = -0.15$, 95% CI [-0.25, -0.05], $p = .001$, indicating partial mediation. These analyses show that gossip to relevant others increased honesty, which was partly driven by reputational concern.

7. Discussion

Dishonesty comes with significant costs to society, and mitigating dishonesty is therefore crucial. Increasing observability has been suggested as an effective remedy against dishonesty, but research findings provided mixed evidence for this effectiveness. We proposed that whether observability increases honesty depends on whether there is an opportunity for gossip to relevant others, because this increases reputational concern. Consistent with this argument, data from 2 incentivized behavioral experiments revealed that people were more honest when they knew observers could gossip about their behavior to relevant others compared to when observers could not gossip or they were not observed at all. Thus, merely being observed by another participant did not increase people’s honesty, but the possibility of being gossiped about to relevant others, that is, those who could provide financial benefits in the future, did. This effect was partly mediated by reputational concern.

Further analyses revealed that participants’ increased reputational concern in the gossip condition is warranted: Participants who reported higher die roll outcomes were more likely to be portrayed as dishonest in gossip messages, and gossip recipients in turn conditioned their trust in a subsequent trust game on the content of the gossip they received (i.e., they trusted those portrayed as dishonest less than

those portrayed as honest). Interestingly, whereas those portrayed as dishonest were trusted less, those portrayed as honest were not trusted more compared to when no information was provided, suggesting that the default presumption is that partners are honest and this changes when gossip informs recipients otherwise. This is in line with earlier findings that gossip can be used to expose norm violators and thereby warn gossip recipients about the target's behavior (Beersma and Van Kleef, 2012; Feinberg et al., 2012), and the Truth Default Theory which proposes that people assume other people's messages to be honest (Levine, 2015), and negative judgment bias (Skowronski and Carlson, 1987).

The current evidence that observability combined with gossip to relevant others increases honesty due to reputational concern helps explain previous conflicting findings regarding the effectiveness of observability in increasing honesty. For example, that observability does not increase honesty in one-shot interactions may be caused by people not being as concerned about their reputation in these settings, in which their reputation in the eyes of others has little consequence. In contrast, in repeated interactions, people are more concerned about their reputation, and therefore observability increases honesty in these settings (Bodenschatz and Irlenbusch, 2019).

In Study 2, we did not find evidence that gossip to *irrelevant* others increases honesty. According to theory about indirect reciprocity (e.g., Wu et al., 2015), gossip increases reputational concerns because the gossip recipient can (potentially) harm or benefit the gossip target in the future or transmit information obtained via gossip to relevant others who in turn can harm or benefit the target's outcomes. In daily life, people have been found to most often gossip to others within social networks (Dores Cruz et al., 2020), and under such circumstances, targets cannot eliminate the possibility that the gossip might be transmitted to someone who can benefit or harm them in the future. In Study 2, we experimentally varied the possibility of gossip being transmitted to relevant versus irrelevant others. We found that if the gossip recipient did not have the opportunity to benefit or harm the gossip target in a trust game (because the trust game was not incentivized), gossip did not increase honesty, attesting to the importance of indirect reciprocity in explaining the effects of gossip. These results indicate that reputational concerns are triggered when people can gossip to relevant others.

7.1. Limitations and suggestions for future research

It is unclear whether and how our results may have been different if we had collected data in an offline (lab) setting. On the one hand, the fear of reputational damage may be stronger in offline settings. This could lead people to behave more honestly across the board, even in the absence of gossip, potentially resulting in smaller effects of gossip and observability. On the other hand, the impact of gossip is likely to be stronger in offline situations, as there may be a greater perceived risk of consequential reputational damage due to lower anonymity. We are inclined to expect the latter, that is, the effects of gossip would be stronger in an offline laboratory setting (and arguably even stronger in real-life settings where usually more is at stake than could be simulated in an experimental context while adhering to ethical guidelines). This would imply that the tests of our hypotheses in an online context constitute conservative tests (Mook, 1983). Nevertheless, whether effects are weaker or stronger in offline settings is an empirical question which we see as an interesting direction for future research.

Another aspect of our studies that may be considered as a limitation is that we, as experimenters, could not observe the participants rolling the actual die rolls. This implies that potentially, participants might have used strategies to justify their lies, for example, by not rolling the die in the first place. Future studies could investigate if (and when) participants are likely to use such strategies. Importantly, even if participants in our studies used such strategies, their behavior still qualifies as dishonest, and is, therefore, a valid operationalization of our dependent variable: whether people are dishonest for their benefit.

Another notable issue is that in Study 2, the correlation between dishonesty and the type of gossip message sent was smaller ($r = .24$) than in Study 1 ($r = .77$). A possible explanation is that participants in Study 1 could gossip using a free-form text message. In contrast, in Study 2, we modified the procedure such that participants could select a standardized gossip statement. The latter made more subtle communication (i.e., communicating that the participant one observed behaved dishonestly by

making a joke about it, or a cynical remark such as: ‘Player A magically rolled a six 30 times in a roll, AMAZING’) difficult, as participants instead had to quite bluntly state what their evaluation of the target’s (dis)honesty was. Importantly, however, the effect of the possibility to send gossip on dishonesty was similar across studies, as was the effect of received gossip on trust. Future studies could investigate differences between free-format and standardized gossip.

Furthermore, in Study 2, we did not find evidence that gossip to irrelevant others impacts people’s reputational concerns and thereby dishonesty. In contrast, when gossip to relevant others was possible, people were more concerned about their reputation and thereby altered their behavior and behaved more honestly. This highlights that reputational concerns are not automatically triggered when people can talk about one’s behavior, rather, it is important that gossip has potential consequences. This finding can be explained by indirect reciprocity (Wu et al., 2016): People care about their reputation because it can have consequences. Future studies could further examine the conditions under which reputational concerns matter.

Lastly, reputational concerns are multi-faceted, that is, people can be concerned about their reputation on various dimensions. For example, one can be concerned about appearing prosocial and about appearing honestly. In our studies, we tested the effects of concerns about one’s reputation to appear honestly. Yet, in a setting in which dishonest behavior might benefit another person (e.g., Leib et al., 2021), people might be concerned just as much, or even more, about their reputation of being a prosocial person. As such, if people prefer to appear prosocial over appearing honest, reputational concerns could in these settings even increase dishonesty. Future studies could investigate these possibilities to further understand the impact of reputational concerns on (dis)honest behavior.

8. Conclusion

We showed that the effectiveness of observability in increasing honesty depends on whether observability increases reputational concern, which it does when observers have an opportunity to gossip to relevant others. This finding enhances understanding of when and how observability increases honesty and illuminates how a ubiquitous activity in everyday life—gossip—can be wielded for the good.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/jdm.2024.10>.

References

- Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, 87(4), 1115–1153. <https://doi.org/10.3982/ECTA14673>
- Bäker, A., & Mechtel, M. (2019). The impact of peer presence on cheating. *Economic Inquiry*, 57(2), 792–812.
- Balliet, D., & Van Lange, P. A. M. (2013). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin*, 139(5), 1090–1112. <https://doi.org/10.1037/a0030939>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2019). *lme4: Linear mixed-effects models using Eigen and S4*. <http://CRAN.R-project.org/package=lme4>
- Beersma, B., & Van Kleef, G. A. (2011). How the grapevine keeps you in line: Gossip increases contributions to the group. *Social Psychological and Personality Science*, 2(6), 642–649. <https://doi.org/10.1177/1948550611405073>
- Beersma, B., & Van Kleef, G. A. (2012). Why people gossip: An empirical analysis of social motives, antecedents, and consequences: Why people gossip. *Journal of Applied Social Psychology*, 42(11), 2640–2670. <https://doi.org/10.1111/j.1559-1816.2012.00956.x>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142.
- Bø, E. E., Slemrod, J., & Thoresen, T. O. (2015). Taxes on the internet: Deterrence effects of public disclosure. *American Economic Journal: Economic Policy*, 7(1), 36–62. <https://doi.org/10.1257/pol.20130330>
- Bodenschatz, A., & Irlenbusch, B. (2019). Do two bribe less than one? – An experimental study on the four-eyes-principle. *Applied Economics Letters*, 26(3), 191–195. <https://doi.org/10.1080/13504851.2018.1456644>
- Caldwell, R. L. (1986). The deceptive use of reputation by stomatopods. In *Deception: Perspectives on human and nonhuman deceit* (pp. 129–145). New York: State University of New York Press.
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon’s MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>

- Dores Cruz, T. D., Thielmann, I., Columbus, S., Molho, C., Wu, J., Righetti, F., de Vries, R. E., Koutsoumpis, A., van Lange, P., Beersma, B., & Balliet, D. (2020). *Gossip and reputation in everyday life* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/y79qk>
- Dores Cruz, T. D., Nieper, A. S., Testori, M., Martinescu, E., & Beersma, B. (2021). An integrative definition and framework to study gossip. *Group & Organization Management*, *46*(2), 252–285.
- Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, *20*, 92–95. <https://doi.org/10.1016/j.copsyc.2017.08.024>
- Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: Reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology*, *102*(5), 1015–1030. <https://doi.org/10.1037/a0026650>
- Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise: An experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547. <https://doi.org/10.1111/jeea.12014>
- Fonseca, M. A., & Peters, K. (2018). Will any gossip do? Gossip does not need to be perfectly accurate to promote trust. *Games and Economic Behavior*, *107*, 253–281. <https://doi.org/10.1016/j.geb.2017.09.015>
- Hendriks, A. (2012). *SoPHIE: Software platform for human interaction experiments*. Osnabrück: University of Osnabrueck.
- Houser, D., List, J. A., Piovesan, M., Samek, A., & Winter, J. (2016). Dishonesty: From parents to children. *European Economic Review*, *82*, 242–254. <https://doi.org/10.1016/j.euroecorev.2015.11.003>
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*(5), 865–889. <https://doi.org/10.1016/j.joep.2011.05.007>
- Köbis, N., van der Lingen, S., Dores Cruz, T. D., Iragorri-Carter, D., van Prooijen, J.-W., Righetti, F., & van Lange, P. (2019). *The look over your shoulder: Unethical behaviour decreases in the physical presence of observers* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/gxu96>
- Lakens, D. (2016). *One-sided tests: Efficient and underused*. <http://daniellakens.blogspot.com/2016/03/one-sided-tests-efficient-and-underused.html>
- Leib, M., Köbis, N., Soraperra, I., Weisel, O., & Shalvi, S. (2021). Collaborative dishonesty: A meta-analytic review. *Psychological Bulletin*, *147*(12), 1241.
- Lilleholt, L., Schild, C., & Zettler, I. (2020). Not all computerized cheating tasks are equal: A comparison of computerized and non-computerized versions of a cheating task. *Journal of Economic Psychology*, *78*, 102270. <https://doi.org/10.1016/j.joep.2020.102270>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276–282.
- Mook, D. G. (1983). In defense of external invalidity. *American psychologist*, *38*(4), 379.
- Pagliaro, S., Brambilla, M., Sacchi, S., D'Angelo, M., & Ellemers, N. (2013). Initial impressions determine behaviours: Morality predicts the willingness to help newcomers. *Journal of Business Ethics*, *117*(1), 37–44. <https://doi.org/10.1007/s10551-012-1508-y>
- Peters, K., & Fonseca, M. A. (2020). Truth, lies, and gossip. *Psychological Science*, *31*(6), 702–714. <https://doi.org/10.1177/0956797620916708>
- Pfattheicher, S., Schindler, S., & Nockur, L. (2019). On the impact of Honesty-Humility and a cue of being watched on cheating behavior. *Journal of Economic Psychology*, *71*, 159–174.
- Robbins, M. L., & Karan, A. (2020). Who gossips and how in everyday life? *Social Psychological and Personality Science*, *11*(2), 185–195. <https://doi.org/10.1177/1948550619837000>
- Slemrod, J. B. (2016). *Tax compliance and enforcement: New research and its policy implications*. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.2726077>
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38. <https://doi.org/10.18637/jss.v059.i05>
- Weisel, O., & Shalvi, S. (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences*, *112*(34), 10651–10656. <https://doi.org/10.1073/pnas.1423035112>
- West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of evolutionary biology*, *20*(2), 415–432.
- Wu, J., Balliet, D., & Van Lange, P. A. M. (2015). When does gossip promote generosity? Indirect reciprocity under the shadow of the future. *Social Psychological and Personality Science*, *6*(8), 923–930. <https://doi.org/10.1177/1948550615595272>

Cite this article: Nieper, A. S., Beersma, B., Dijkstra, M. T. M., and Van Kleef, G. A. (2025). When and why does observability increase honesty? The role of gossip and reputational concern *Judgment and Decision Making*, e5. <https://doi.org/10.1017/jdm.2024.10>