nature human behaviour

Article

https://doi.org/10.1038/s41562-022-01457-w

Corrupt third parties undermine trust and prosocial behaviour between people

Received: 4 August 2021

Accepted: 5 September 2022

Published online: 27 October 2022

Check for updates

Giuliana Spadaro[1,2], Catherine Molho[3,4], Jan-Willem Van Prooijen[1,2,5,6], Angelo Romano[7], Cristina O. Mosso[8] and Paul A. M. Van Lange[1,2]

Corruption is a pervasive phenomenon that affects the quality of institutions, undermines economic growth and exacerbates inequalities around the globe. Here we tested whether perceiving representatives of institutions as corrupt undermines trust and subsequent prosocial behaviour among strangers. We developed an experimental game paradigm modelling representatives as third-party punishers to manipulate or assess corruption and examine its relationship with trust and prosociality (trust behaviour, cooperation and generosity). In a sequential dyadic die-rolling task, the participants observed the dishonest behaviour of a target who would subsequently serve as a third-party punisher in a trust game (Study 1a, $N = 540$), in a prisoner's dilemma (Study 1b, $N = 503$) and in dictator games (Studies 2–4, $N = 765$, pre-registered). Across these five studies, perceiving a third party as corrupt undermined interpersonal trust and, in turn, prosocial behaviour. These findings contribute to our understanding of the critical role that representatives of institutions play in shaping cooperative relationships in modern societies.

In 2015, an anonymous source leaked 11.5 million documents from the fourth-biggest offshore law firm in the world. This leak unveiled a system of rich people, politicians, public officials and close associates who exploited their privileged positions to engage in tax evasion, fraud and evasion of international sanctions. A total of 600 people from 42 countries were involved in what is considered one of the biggest leaks ever reported[1]. This scandal, now known as the Panama papers, pointed at a long-standing problem of institutional representatives taking advantage of their entrusted positions to gain private benefits[2,3]. The institutional challenges that such scandals pose have been extensively examined[4], but less attention has been devoted to the question of whether knowing about the dishonesty and corruption of institutional representatives affects interpersonal trust and prosocial behaviour towards fellow citizens.

Corruption is a critical societal and scientific issue that has attracted considerable research interest across many disciplines, such as economics, political science, sociology, law and psychology[5–8]. Decades of research have aimed to understand this phenomenon and its dramatic consequences on societies, as it undermines economic growth and exacerbates inequalities around the globe[5,9,10]. Many studies focused on cross-cultural differences in corruption levels, while other research investigated what makes people engage in corrupt behaviour[7,11,12]. Importantly, it has been hypothesized that corruption may affect social interactions involving interpersonal trust and prosocial behaviour[13–15]. In daily life, people learn about corruption by directly or indirectly witnessing the behaviour of their representatives, such as public officials that accept bribes or politicians that evade taxes, which in turn affects their trust towards institutions[16]. Yet, to date, experimental evidence on whether perceiving corruption by representatives of institutions undermines trust and cooperation towards strangers is lacking.

[1]Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [2]Institute for Brain and Behavior Amsterdam (IBBA), Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [3]Institute for Advanced Study in Toulouse, Toulouse, France. [4]Center for Research in Experimental Economics and Political Decision Making (CREED), University of Amsterdam, Amsterdam, the Netherlands. [5]The Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), Amsterdam, the Netherlands. [6] Department of Criminal Law and Criminology, Maastricht University, Maastricht, the Netherlands. [7]Department of Social, Economic and Organisational Psychology, Leiden University, Leiden, the Netherlands. [8]Department of Psychology, University of Turin, Turin, Italy. e-mail: g.spadaro@vu.nl

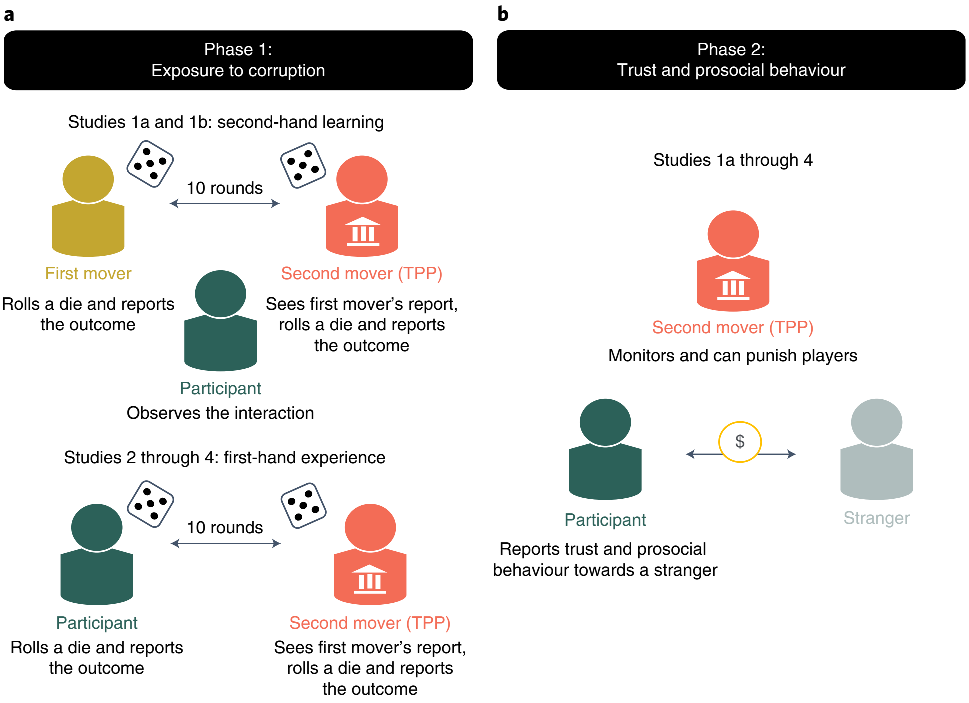
a
Phase 1:
Exposure to corruption
Studies 1a and 1b: second-hand learning
10 rounds
First mover
Rolls a die and reports the outcome
Second mover (TPP)
Sees first mover's report, rolls a die and reports the outcome
Participant
Observes the interaction
Studies 2 through 4: first-hand experience
10 rounds
Participant
Rolls a die and reports the outcome
Second mover (TPP)
Sees first mover's report, rolls a die and reports the outcome
b
Phase 2:
Trust and prosocial behaviour
Studies 1a through 4
Second mover (TPP)
Monitors and can punish players
$
Participant
Reports trust and prosocial behaviour towards a stranger
Stranger

Fig. 1 | The two phases of the experimental procedure of Studies 1a through 4. a, Participants are exposed to corruption of the TPP by either observing the interaction between two players in the die-rolling (Studies 1a and 1b) or personally interacting in the die-rolling task as first movers (Studies 2 through 4). **b**, Participants report their trust and prosocial behaviour towards an unrelated stranger in an economic game where the second mover serves as a TPP.

In the interdisciplinary literature on corruption and trust, two major streams of research have been advanced. One may be labelled the bottom-up perspective, which assumes that the effectiveness of institutions depends largely on informal social processes, such as individuals' ability to solve local and small-scale social dilemmas[17–19]. In this view, interpersonal trust is considered the basis for ensuring the effectiveness of institutions[20]. A second perspective, which may be referred to as top-down, assumes that institutions shape human interactions and therefore influence interpersonal trust and cooperation. Here interpersonal trust is considered a result of the quality of institutions. Often, this perspective goes even further by suggesting that one of the main functions of institutions is to mitigate vulnerability in interactions with strangers[21,22]. If public institutions and their representatives are perceived as unable to provide security, then interpersonal trust can emerge only in narrow and tight networks. Yet, in modern globalized contexts, transactions with strangers are frequent and necessitate building generalized trust[23].

There is some empirical evidence that provides support for this top-down hypothesis, showing that interpersonal trust increases among individuals who migrate to countries with lower levels of corruption[24] and that institutional trust is one of the strongest predictors of interpersonal trust[25]. Notably, research suggests that experiencing corruption enacted by public officials or by other strangers is associated with individual behaviours, such as honesty and ingroup solidarity. In fact, individuals display less ethical values when they are exposed to institutions with more prevalent corrupt practices[26,27]. Moreover, the mere observation of corrupt behaviour enacted by neighbours or ingroup members seems associated with individuals' propensity to act dishonestly[11,28]. Yet, the relation between corruption and generalized trust is still an unsolved issue. Importantly, if corrupt representatives of institutions have a negative effect on trust, this may also have crucial implications for prosocial behaviour between strangers. Indeed, trust is one of the most influential factors that determine cooperation in situations when a conflict between individual and collective interests occurs[29]. As the implementation of third-party sanctioning institutions is one of the most powerful strategies to promote prosocial behaviour in the absence of reputational information[30], it becomes crucial to understand whether the corruption of such third parties may undermine the effectiveness of sanctioning[15].

Individuals witness norm violations from peers daily[31], but they are also exposed to violations from representatives of public institutions. Here we examine whether learning that institutional representatives are corrupt (that is, act dishonestly to enhance their self-benefit and use their power to profit at the expense of the collective) undermines trust and cooperation towards strangers. In a set of five studies, we distinguish between two sources of perceived corruption that may underlie beliefs about corrupt institutions in everyday life and negatively affect trust towards institutional representatives: second-hand learning (for example, political scandals broadcasted in media) and first-hand experience (for example, personal experience with corrupt authorities accepting bribes)[32]. Second-hand learning of corruption is very frequent in daily life and has been associated with sudden declines in trust towards political representatives[33,34]. First-hand experience of corruption may be less ubiquitous in some contexts or cultures[26], but it elicits long-lasting negative societal outcomes[15,35].

We developed an experimental paradigm that is rooted in the tradition of research using economic games[36,37]. In this paradigm, individuals can decide to trust and behave prosocially towards others under the scrutiny of a third-party observer that proved to be corrupt (or not) in a previous interaction. Some daily life interactions between strangers indeed occur under the oversight of institutional representatives, but many others might not directly involve representatives of institutions. Yet, to study the effect of corrupt institutional representatives in a controlled experimental setting, we operationalized them as observably corrupt or honest third parties who monitored and regulated economic transactions in an incentivized experiment. The game is divided into two phases (Fig. 1).

In Phase 1, the participants observe a person cheating (or not) in a sequential dyadic die-rolling task—that is, a situation that allows one of the interactants to profit by acting dishonestly[7]. In this task, two players are instructed to roll a six-sided die privately and sequentially and to report the given outcome. The first mover earns a monetary payoff regardless of the outcomes of the die roll, while the second mover receives a payoff only when their outcome exactly matches the declared outcome of the first mover. Importantly, in this task, the second mover knows in advance the reported outcome of the first mover, while it is impossible for the experimenter to verify whether the second mover's declared outcome corresponds with their actual die roll. Hence, for the second mover, this situation captures specific dishonest behaviours that are closely linked to corruption, as the second mover is tempted to misuse the information and declare corresponding outcomes for self-benefit. This paradigm has been demonstrated to elicit dishonest behaviour from participants, whose reports deviate substantially from reports expected by chance[7,38,39]. Such too-good-to-be-true outcomes are unambiguously interpreted by others as dishonest behaviour[40,41].

To resemble real-life situations where people can learn about corruption indirectly or directly, the sequential die-rolling task enables participants to either (1) observe an interaction between the two movers and learn that the second mover behaved honestly or dishonestly (second-hand learning; Studies 1a and 1b), or (2) personally engage in the die-rolling task as first movers and experience the second mover's honesty or dishonesty themselves (first-hand experience; Studies 2 through 4). Additionally, while Studies 1a through 2 specifically focus on dishonest behaviour that enhances self-benefit (a specific feature of corruption), in Studies 3 and 4 we model corrupt behaviour to also include its negative externalities for the collective. Specifically, the second mover's dishonesty in Phase 1 enhances self-benefit and directly harms the collective by allowing the second mover to take possession of resources that would otherwise benefit all participants (Study 3) or the broader collective (Study 4).

The participants then transition to Phase 2 of the game. Here they learn that the person they just observed or interacted with as the second mover will serve as a third-party punisher (TPP) in an economic game (specifically, a trust game (TG), a prisoner's dilemma game (PD) or a dictator game (DG)) where they can decide whether to behave prosocially towards a stranger. The participants are informed that the TPP (previously the second mover they learned to be either honest or corrupt) can invest his or her own endowment to reduce players' outcomes in the game on the basis of their behaviour. This implementation of TPP has been used extensively in previous research using economic games to model the behaviour of institutional representatives[42]. In these settings, it is common to observe punishing behaviour from third parties, even if it is costly and seemingly at odds with self-interest[43]. We measure interpersonal trust using a six-item scale, which asks the participants how much they trust a new partner (that is, a participant who was not part of the die-rolling in Phase 1) with whom they are matched in the one-shot TPP economic games played in Phase 2 (ref. [44]). The use of this measure has the additional advantage of allowing us to zero in on a potentially key mechanism underlying the effects of corrupt institutions on prosocial behaviour (and disentangle it from other explanations—for example, those based on beliefs about the corrupt third parties' punishment behaviour). We measure different forms of prosocial behaviour (namely trust behaviour, cooperation and generosity) as the amount given to this new partner in Phase 2 in multiple economic games[37,45].

Our main question here is whether knowing that a TPP has behaved dishonestly or honestly in the past affects interpersonal trust and prosocial behaviour towards an unrelated stranger. Across five studies, we tested two hypotheses. The first hypothesis is that observing corrupt behaviour by a person serving as a third party who administers sanctions will undermine trust towards a stranger. The second hypothesis is that the influence of corruption on interpersonal trust should, in turn, undermine prosocial behaviour towards the same unknown partner in an economic game.

Studies 1a and 1b provided a preliminary, internally valid test for our hypotheses and focused on second-hand learning of dishonest behaviour displayed by a third party with punishment capacity. The participants observed an ostensible die-rolling task interaction between two movers in Phase 1. We manipulated corruption through varying the degree of cheating of the second mover by providing pre-programmed feedback about both players' behaviour in the dyadic die-rolling task (one out of ten reported doubles versus ten out of ten reported doubles). The second mover then served as a TPP in Phase 2. Additionally, we included a control condition in which the participants observed a player reporting ten out of ten doubles but not acting as the TPP in Phase 2. In this condition, the TPP was a stranger about whom the participants had no reputational information. We then assessed self-reported interpersonal trust and trust behaviour towards an unknown partner in a TG involving third-party punishment[46]. In Study 1b, we replicated the design of Study 1a but examined interpersonal trust and cooperation in a TPP PD[47].

Study 2 provided a pre-registered replication of our findings and tested the hypotheses in an observational setting in which the participants could directly interact with a potentially corrupt (or honest) future TPP. Contrary to Studies 1a and 1b, the participants could observe naturally emerging levels of corruption from the third party—a feature that may better align with experiences from everyday life. Specifically, in Phase 1, the participants were paired to take part in the sequential dyadic die-rolling task[7]. They then transitioned to Phase 2 of the game, in which we measured the first movers' interpersonal trust towards a stranger and generosity in a TPP DG. The participants previously acting as second movers acted as the TPP in Phase 2.

Studies 3 and 4 expanded on the previous studies by providing a pre-registered test of our hypotheses in a more ecologically valid observational setting, which focused on first-hand experience of a third party misusing their power to profit at the expense of the collective. To do so, we adapted the procedure employed in Study 2 by introducing a different incentive structure that more closely resembles the definition of corruption as 'abuse of public means for private gain'[2] and 'power asymmetry over shared resources'[3]. Accordingly, dishonest behaviour from the second movers in the die-rolling task (and the subsequent TPP in the DG) directly resulted in the depletion of a common good relevant to the community of participants involved in the study (a fund to be equally allocated among all participants in the study; Study 3) or to the broader human collective (a fund to be allocated to a pro-environmental charity; Study 4).

Results

Analyses of all studies were conducted in R (v4.0.5)[48] using linear models and the PROCESS macro for mediation[49]. Statistical assumptions were formally tested. Their full report, along with robustness checks in case of violation, are reported in the Supplementary Information ('Formal test of assumptions' section).

Study 1a

To test whether the manipulation of corruption was successful, we asked the participants to what extent they perceived the second mover as honest in reporting his or her score on a seven-point Likert scale (1 = completely dishonest, 7 = completely honest). We reverse-scored this item for easier interpretation, with high scores indicating greater perceived dishonesty. The manipulation resulted in greater perceptions of dishonesty when second movers reported 10/10 doubles (mean = 5.60, s.d. = 1.91) compared with 1/10 doubles (mean = 1.55, s.d. = 1.08) ($t(538) = -26.43$; $P < 0.001$; effect size, Cohen's $d = 2.41$; 95% confidence interval (CI) (2.11, 2.72)). A one-way analysis of variance testing the effects of corruption of the TPP on self-reported interpersonal trust towards the stranger in the TG revealed a main effect of the manipulation of corruption ($F(2, 537) = 4.67$; $P = 0.010$; effect size, $\eta^2_{\mathrm{p}} = 0.02$; 95% CI (0.01, 0.04)) (Supplementary Table 1). We created two hypotheses-relevant orthogonal contrasts of our experimental conditions: Contrast 1 (corrupt TPP versus honest TPP and control conditions) and Contrast 2 (honest TPP versus control conditions). Planned comparisons revealed a significant Contrast 1 ($F(1, 538) = 5.63$; $P = 0.018$; $d = 0.22$; 95% CI (0.04, 0.40)), indicating less self-reported trust towards a stranger when being monitored by a corrupt TPP (mean = 4.50, s.d. = 1.36), compared with an honest TPP and an unknown TPP (mean = 4.80, s.d. = 1.41). Although the pattern of results related to Contrast 2 suggests that mere exposure to corruption might undermine interpersonal trust, we found no credible evidence for differences in interpersonal trust between the honest TPP and the control condition ($F(1, 358) = 3.62$; $P = 0.058$; $\eta^2_p = 0.01$; 95% CI (0.01, 0.04)). We thus find little evidence of a difference in interpersonal trust between the honest TPP and the control condition. We then tested whether knowing about the corruption of third parties (the corrupt TPP condition) affected self-reported trust and, in turn, trust behaviour using the bootstrapping method for mediation analysis[49]. The results show evidence of a significant indirect effect of corruption on trust behaviour in the TG via interpersonal trust (unstandardized regression coefficient, $b = 0.21$; 95% CI (0.07, 0.37)). Hence, our first study provides initial evidence that perceiving institutional representatives as corrupt undermines trust towards strangers and in turn reduces prosocial behaviour.

Study 1b

In Study 1b, the manipulation was again successful in affecting the perceived dishonesty of the TPP ($t(501) = 21.93$; $P < 0.001$; $d = -1.96$; 95% CI (−2.20, −1.71)), with the TPP being perceived as more dishonest in the corrupt TPP condition (mean = 5.32, s.d. = 1.98) than in the honest TPP condition (mean = 1.93, s.d. = 1.46). The results of Study 1b reveal that participants who faced a corrupt TPP trusted their partner less (mean = 4.91, s.d. = 1.42) than participants who faced an honest TPP (mean = 5.22, s.d. = 1.25) ($F(1, 501) = 6.80$; $P = 0.009$; $d = 0.23$; 95% CI (0.06, 0.41)). We then tested whether perceiving the third party as corrupt affected cooperation in the PD indirectly through interpersonal trust, using the bootstrapping method for mediation analysis[49]. The results show a significant indirect effect of corruption on cooperation via interpersonal trust ($b = 1.82$; 95% CI (0.40, 3.44)). Altogether, these results replicate the findings of Study 1a, showing a negative effect of corruption on trust and subsequent prosocial behaviour.

Study 2

We conducted a simple linear regression in which interpersonal trust towards the stranger in the DG was regressed on the sender's perceptions of the dishonesty of the TPP in reporting the outcomes in the die-rolling task. Consistent with our hypothesis, the perceived dishonesty of the TPP was significantly and negatively associated with the extent to which the senders trusted the receivers (standardized regression coefficient, $\beta = -0.39$; $t(189) = -5.88$; $P < 0.001$; 95% CI (−0.52, −0.26)) and explained a significant proportion of variance (effect size, $R^2 = 0.15$, $F(1, 189) = 34.55$, $P < 0.001$) (Fig. 2). We then tested whether interacting with a TPP (perceived as honest or dishonest) in a previous die-rolling task would be indirectly associated with generosity via interpersonal trust. Using the bootstrapping method for mediation[49], we replicated the findings of the previous studies, showing a significant indirect effect on generosity via interpersonal trust ($b = -0.63$; 95% CI (−1.25, −0.06)). Overall, Study 2 presents compelling evidence in a real interaction setting that the more the participants perceived the second movers (and subsequent TPPs) as corrupt in the die-rolling task, the less they trusted an unrelated player in the subsequent DG with third-party sanctioning. Moreover, we found again that this decline in trust was negatively associated with prosocial behaviour.

Studies 3 and 4

Studies 3 and 4 used an incentive structure in which corruption was operationalized in terms of self-benefit and detrimental consequences for the collective. Again, the findings of both studies supported both pre-registered hypotheses, thereby fully replicating the patterns observed in Study 2. In both studies, the perceived dishonesty of the third party was significantly and negatively associated with the extent to which the senders trusted the receivers in the DG ($R^2_{\text{Study 3}} = 0.19$, $R^2_{\text{Study 4}} = 0.11$) (Fig. 2) and showed a significant indirect effect on generosity via interpersonal trust ($b = -2.47$; 95% $\text{CI}_{\text{Study 3}}$, (−5.07, −0.66); $b = -2.04$; 95% $\text{CI}_{\text{Study 4}}$, (−3.79, −0.58)). Notably, the findings of Study 4 remained significant when we controlled for the sender's subjective importance of the mission of the charity. The results are presented in detail in the Supplementary Information, along with robustness checks (sections 'Study 3' and 'Study 4'). An internal random effects meta-analysis of the correlation between the perception of the TPP's dishonesty in the die-rolling task and interpersonal trust towards the stranger in Studies 2–4 displayed a medium-size negative meta-analytic correlation ($k = 3$ effects; $r = -0.40$; $P < 0.001$; 95% CI (−0.48, −0.32)).

Across the five studies, we found that interpersonal trust consistently mediated the relationship between corruption and prosocial behaviour. In Studies 3 and 4, we additionally included a measure of the senders' expectations about punishment enacted by the TPP in the DG to explore another potential underlying mechanism. Specifically, we aimed to explore whether the changes in prosocial behaviour after interacting with a corrupt TPP were attributable to the belief that this TPP would (not) punish selfish senders. Given that punishment decisions are costly, dishonest third parties could be expected to be more selfish and punish less, and this might explain why senders were less generous towards others in the presence of a dishonest TPP. The results of the two studies provided inconsistent findings, revealing an indirect effect via punishment expectations on generosity in Study 3 (95% CI (−5.99, −2.24)), but no evidence for this indirect effect in Study 4 (95% CI (−2.21, 1.88)) (Supplementary Information, sections 'Study 3' and 'Study 4'). Our exploratory findings thus do not provide enough evidence to support (or rule out) this potential additional mechanism.

Discussion

Considerable research in various scientific disciplines has addressed the intricate associations between the degree to which institutions are corrupt and the extent to which people trust one another and build cooperative relations. One perspective suggests that the success of institutions is rooted in interpersonal processes such as trust[18]. Another perspective assumes a top-down process, suggesting that the functioning of institutions serves as a basis to promote and sustain interpersonal trust[22,25]. However, as far as we know, this latter claim has not been tested in experimental settings.

In the present research, we provided an initial test of a top-down perspective, examining the role of a corrupt versus honest institutional representative, here operationalized as a third-party observer with the power to regulate interaction through punishment. To do so, we revisited the sequential dyadic die-rolling paradigm where the participants could learn whether the third party was corrupt or not via second-hand learning or via first-hand experience. Across five studies ($N = 1{,}808$), we found support for the central hypothesis guiding this research: perceiving third parties as corrupt is associated with a decline in interpersonal trust, and subsequent prosocial behaviour, towards strangers. This result was robust across a broad set of economic games and designs. Our findings contribute to the trust literature by suggesting that institutional representatives exert substantial influence on interpersonal trust within societies. Hence, this can be interpreted as initial evidence for a top-down route in the relation between institutions, trust and prosocial behaviour. Importantly, this evidence does not rule out the influence of bottom-up processes, which could coexist and even have a reciprocal influence[50]. However, it is almost inevitable that at least some dishonest behaviour by institutions will eventually reach the public eye, which in turn endangers interpersonal trust. The result of such a challenge to trust can be dramatic, because the repair of trust may not always last in the long term, and it is a process that requires considerable effort and

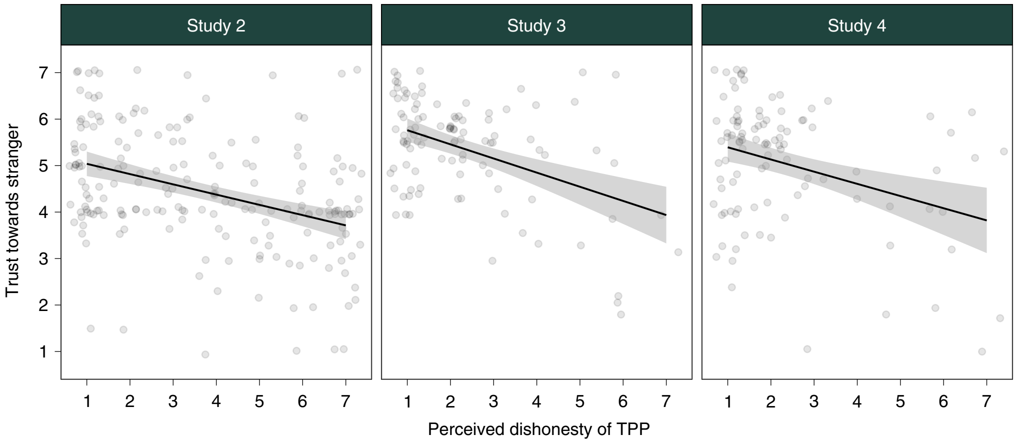

Fig. 2 | Relationship between the perceived dishonesty of the TPP and interpersonal trust across Studies 2 through 4. Each graph was obtained by regressing the first movers' self-reported trust towards the recipient in the DG on their self-reported perceived dishonesty of the TPP in the sequential dyadic die-rolling task. Each dot represents an observation. The shaded area indicates the 95% CI. Across the three studies, the perceived dishonesty of the TPP was significantly and negatively associated with the extent to which the senders trusted the receivers: in Study 2 ($k = 191$ dyads), $\beta = -0.39$; $t(189) = -5.88$; $P < 0.001$; 95% CI (−0.52, −0.26) (two-sided); $R^2 = 0.15$; in Study 3 ($k = 102$ dyads), $\beta = -0.44$; $t(100) = -4.947$; $P < 0.001$; 95% CI (−0.62, −0.27) (one-sided); $R^2 = 0.19$; and in Study 4 ($k = 101$ dyads), $\beta = -0.35$; $t(99) = -3.71$; $P < 0.001$; 95% CI (−0.53, −0.16) (one-sided); $R^2 = 0.11$.

time[51,52]. Because trust is essential to well-functioning groups, organizations and societies[50], a lack of trust can at least temporarily undermine societal development, as it is related to important societal outcomes such as economic growth and political participation[53,54].

We also found that lower levels of interpersonal trust associated with corrupt institutional representatives were, in turn, related to a decrease in prosocial behaviour. Promoting cooperation and prosociality towards strangers is essential to solving important problems such as global warming, pollution, tax evasion and other societal collective challenges[29,55]. If such an effect of corruption on prosocial behaviour occurs, future interventions should be implemented following a top-down approach that starts from institutional representatives, rather than horizontally between individuals[51]. If citizens tend to distrust each other as a result of knowing that institutional representatives are corrupt, the implementation of punishment or reputational systems may not be effective or may even backfire, crowding out interpersonal and institutional trust or giving rise to antisocial punishment[17,56].

Before closing, we briefly discuss some limitations and avenues for future research. First, in the present research, institutional representatives were operationalized at the micro level as a third-party sanctioning actor in cooperative exchanges. Although this operationalization is commonly used in the experimental literature[57,58], it does not fully capture the complex and encompassing world of institutions that most people experience in everyday life. Second, the online setting also differs from daily experiences with institutional representatives. One key difference is that these experiences are often repeated (rather than one-shot) and usually extend over substantial periods[15]. Therefore, in everyday life people may often come to internalize norms of corruption, and the detrimental effects of corruption may be even more dramatic and enduring. This seems especially true for some countries in which individuals regularly observe and need to interact with corrupt representatives with sanctioning power[13].

Moreover, in many daily situations, people decide to trust and act prosocially towards strangers without any institutional representatives being present or directly involved in the exchange. In such situations, people may act on their internalized beliefs about whether institutions are (or are not) corrupt, built on repeated previous experiences[25], therefore negatively affecting their trust and prosocial behaviour. Such beliefs may vary across different types of institutions (for example, political or legal) and across societies. Indeed, previous research has suggested that when individuals migrate to less corrupt societies, their levels of interpersonal trust change accordingly[24]. For our research questions, it was crucial to specifically determine an association between corrupt representatives of institutions and prosocial behaviour (and to disentangle it from a mere exposure to corrupt behaviour by an unrelated participant), but future research can investigate the detrimental effect of institutions, even when no third parties are actually monitoring interactions between strangers. For that reason, Study 1a included a control condition in which the participants observed dishonest behaviour but then were matched with an unrelated third party. Although there was no credible evidence for the difference between the control and honest TPP conditions ($P = 0.058$), the pattern of results does not exclude the possibility that mere exposure to corruption can also be negatively associated with trust. This would also be consistent with general possibilities discussed in the literature[59,60], even though we do not know of clear-cut evidence demonstrating such effects for observing only one person's prosocial behaviour[61]. Future research can investigate this possibility, as interactions with corrupt peers can occur frequently in daily life. Last, it is worth noting that in Studies 2 through 4, institutional corruption was not experimentally manipulated, and this might limit our ability to make causal claims. That said, the pattern of findings is in line with causal evidence from Studies 1a and 1b that involved an experimental manipulation.

Our current findings represent the beginning of a line of research aimed at identifying the effects of corrupt representatives of institutions on trust and prosocial behaviour. In this first set of results, we provide robust evidence about the negative association between corrupt institutions, trust and prosocial behaviour. Future research is needed to provide further evidence to support (or rule out) specific additional mechanisms underlying the complex relationship between corruption and prosocial behaviour. Here we started by examining the roles of interpersonal trust and beliefs about third-party punishment. For example, although it is possible that participants are less generous because corrupt third parties are less likely to enact costly punishment, we found little evidence for this possibility. First, we found no evidence that actual dishonesty was related to beliefs about third-party punishment in Studies 3 and 4, or that perceived dishonesty was related to beliefs about third-party punishment only in Study 3. Together, these results suggest that participants who observe dishonest behaviour do not directly form the expectation that the future third party will punish less in the subsequent task. Second, as the results of Studies 3 and 4 show, beliefs about third-party punishment did not consistently mediate the relation between perceived dishonesty and prosocial behaviour. By contrast, across all five conducted studies, interpersonal trust mediated the relationship between corruption and prosocial behaviour, suggesting that trust is a key mechanism in the relation between institutions and prosociality.

Testing further underlying mechanisms constitutes an important direction for future investigation, a recommendation that we also make for enhancing ecological validity. For example, future research could complement this set of studies by investigating the effect of common—and subtle—cues of dishonest behaviour that characterize real-world trust (for example, facial expressions[62]). Moreover, in our design, the participants did not benefit from the corrupt behaviour of the third party, while in many real-life situations, individuals often directly benefit from corrupt transactions[7]. The incentive structure of the die-rolling paradigm can be flexibly adapted to model this or different corruption dynamics. Future research can thus use this paradigm to examine the boundary conditions of the relationship between corruption, trust and prosocial behaviour in situations where participants benefit from the corrupt transaction.

To conclude, our studies revealed that perceiving institutional representatives as corrupt can undermine (Studies 1a and 1b) and be negatively associated with (Studies 2 to 4) trust and prosocial behaviour. These findings illuminate the vital functions that institutions might have in human psychology, as well as their role in our perception of and behaviour with strangers. Hence, corruption among institutional representatives may facilitate a culture in which corrupt activities not only come to be viewed as relatively common and normative[26] but also give rise to distrust among strangers. The fact that corruption and distrust are partially rooted in institutional representatives is also relevant for policies that focus on reducing corruption in a sustainable manner. One broader implication is that groups and societies should do all they can to attract institutional leaders with integrity and, perhaps equally important, to shape and nurture an environment in which such leaders are less likely to push or cross ethical boundaries.

Methods

The research was approved by the Scientific and Ethical Review Board of the Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Application VCWE-2017-085. The materials and pre-registrations are accessible at https://osf.io/xhq6e (materials for Studies 1a–4) and https://osf.io/sqp2m/registrations (pre-registrations). In all studies, the participants provided their informed consent, and participation was restricted to the United States. The order of all administered tasks and measures was the same for all participants. Data from all participants completing the study were included in the analyses. Unless otherwise specified, all tests reported in the results and a priori power analyses are two-sided.

Study 1a

Participants. An a priori power analysis (G*Power[63]) revealed a required sample size of 528 to achieve a statistical power of 0.80 to detect an effect size of $d = 0.30$ of the corruption manipulation on interpersonal trust. The participants ($N = 540$; 47% women; mean age, 35.44; s.d., 11.10) were recruited from Amazon Mechanical Turk (MTurk) and completed the online study on the platform Qualtrics for US$1. Moreover, they could receive a US$10 lottery prize as an incentive for attention in the die-rolling task and learned that they could earn up to US$1.50 depending on decisions in the TG. Samples from MTurk are heterogeneous in terms of socio-economic and ethnic diversity, and MTurk is a reliable platform on which to perform behavioural tasks[64,65]. We used a between-participants design, in which the participants were randomly assigned to three conditions: honest TPP, corrupt TPP and control.

Die-rolling task (Phase 1). The manipulation of corruption occurred in Phase 1 of the game. In the honest TPP condition, the participants observed a targeted prospective TPP behaving honestly across ten rounds of the sequential dyadic die-rolling task[7]. Specifically, they learned that the second mover (the prospective TPP) mimicked the outcome of the first mover in only one of ten rounds (rounding down the expected number of doubles assuming honest reporting: 1.66). In the corrupt TPP condition, the participants observed the future TPP reporting the same outcome in the die-rolling task on ten out of ten rounds. Importantly, we included a control condition where the participants observed a dishonest player, but during the following game they faced a TPP they had never encountered before. This condition allowed for a preliminary test of whether mere exposure to corrupt behaviour from another participant influences interpersonal trust, rather than the perception of an institutional representative (that is, the TPP), in particular, as corrupt. Across conditions, the participants were not aware that the second mover in Phase 1 would take part in Phase 2. To elicit and incentivize the attention of the participants when they observed behaviour in the die-rolling task, they were informed that they would be eligible for a lottery prize of US$10 in case of all correct answers in the attention check questions. To this purpose, the observers received questions about the rules of the game, the role of each player and the outcomes of the ten rounds. Prior to receiving instructions, the participants were asked to roll a computerized die on an online website in order to increase the belief that the game and the partners were actually interacting. In reality, their reports were pre-programmed feedback provided according to our experimental manipulations.

Economic game with a TPP (Phase 2). In Phase 2, the participants were matched with a stranger and played a TG[46] with third-party punishment. In this game, the participants were endowed with five monetary units (each worth US$0.10) that they could decide to give to the unrelated stranger (the trustee) (0–5). They knew that each monetary unit they sent to the stranger would be tripled, and then the stranger could decide to return (or not) any amount. Importantly, they knew that their decisions would be observed by the TPP, who could then decide to invest (or not invest) part of the endowment to deduct any monetary units that the participant and the trustee earned during the TG. Our dependent measure of interpersonal trust was an adaptation of the general trust scale, a six-item, seven-point Likert scale (1 = strongly disagree, 7 = strongly agree) (example item: 'I believe that Player 2 is trustworthy'; $\alpha = 0.96$)[44]. Higher scores on this scale mean that the participants trusted their partner more.

Study 1b

Participants. The participants ($N = 503$; 49% women; mean age, 34.88 years; s.d., 10.37) were recruited from MTurk and completed the study on the platform Qualtrics for US$1. In addition, the participants had a chance to receive a lottery prize of US$10 as an incentive for attention in the die-rolling task, and a chance to win a US$2 prize depending on decisions in the PD. We used a 2 (corruption: honest versus corrupt TPP) × 2 (communication: present versus absent) between-participants design in which the participants were randomly assigned to the experimental conditions.

Die-rolling task (Phase 1). The die-rolling task was identical to Study 1a. Differently from Study 1a, we manipulated the possibility of receiving a message from the partner prior to the decision in the PD to test whether the negative effects of corrupt institutions hold when a possibility for communication is present (versus absent). The results from these treatments are presented in the Supplementary Information, section 'Study 1b'.

Economic game with a TPP (Phase 2). This phase was identical to Study 1a, except for the use of a PD with third-party punishment to assess cooperative behaviour. Participants previously playing as first movers were endowed with 100 lottery tickets and decided how many to allocate to an unknown partner who would simultaneously make the same decision (0–100), knowing that the amount would be doubled. Participants previously playing as second movers (TPPs) decided how much to invest to reduce others' final earnings. Each lottery ticket accumulated in the PD increased the chance to be awarded a US$2 prize.

Study 2

Participants. An a priori power analysis (G*Power[63]) revealed a required sample size of 380 (190 dyads) to achieve a statistical power of 0.95 to detect an effect size of d = 0.24. The participants (N = 382; 45.5% women; mean age, 37.73 years; s.d., 10.82) were recruited through MTurk and completed the study for US$2.50. Additionally, they could earn an extra bonus in the die-rolling task (up to US$0.60) and could win a US$2 prize depending on decisions in the DG. We conducted the study through the platform SoPHIE[66], which enables real-time interactions among online participants.

Die-rolling task (Phase 1). Once logged into the platform, all participants were randomly paired and assigned to the role of either first movers or second movers. They were then informed about their role in the game and received detailed instructions for the die-rolling task (see 'Study 1a' for the general procedure of the game). To ensure that second movers might engage in dishonest behaviour, we instructed them to either keep an actual die at hand while participating in the study or open a suggested external web page that allowed them to roll a fair six-sided die. The payoff scheme was disclosed to the participants before the game, as in the previous studies. While first movers earned a fixed amount of US$0.20 irrespective of scoring a double in each round, second movers could get triple that amount (US$0.60) if they reported the same outcome of the die roll as the first mover. This removed any incentive for first movers to lie about their outcome, ruling out the possibility of engaging in corrupt cooperation and taking advantage of the eventual dishonesty of the second movers. After each round, both players received real-time feedback on the reported outcomes of the die rolls. At the end of the die-rolling task, we asked the first movers to what extent they perceived the second mover as honest in reporting his or her score on a seven-point Likert scale (1 = completely dishonest, 7 = completely honest) and then reverse-scored for easier interpretation of analyses. This constituted our independent variable.

Economic game with a TPP (Phase 2). The participants then engaged in a DG with third-party punishment. Participants who were previously playing as first movers (senders) were endowed with 100 lottery tickets and decided how much to give to an unknown receiver (0–100), while participants previously playing as second movers (TPPs) decided how much to invest to reduce others' final earnings. Each lottery ticket accumulated in the DG increased the chance of being awarded a US$2 prize. Finally, we assessed interpersonal trust (α = 0.97) as in the previous studies.

Studies 3 and 4

Participants. An a priori sensitivity power analysis (G*Power[63]) revealed that a sample size of 100 dyads would give us a statistical power of 0.80 to detect an effect size of r = 0.24 (one-sided, following the pre-registered unidirectional hypotheses). The participants in Study 3 (N = 197; 35% women; mean age, 35.56 years; s.d., 10.13) and Study 4 (N = 186; 34% women; mean age, 37.49 years; s.d., 10.90) were recruited through MTurk and completed a real-time interaction study for US$2.50 in the platform SoPHIE[66]. Additionally, they could earn an extra bonus in the die-rolling task (up to US$0.60) and could participate in a lottery to win a US$2 prize. In a limited number of experimental sessions (7% in Study 3 and 8% in Study 4), the participants were matched with the experimenter, who would then play as the TPP. Such sessions are included in the current analyses. The results of analyses excluding such sessions were consistent in terms of both the relationship between perceived dishonesty and interpersonal trust and the indirect effect on generosity via trust (Supplementary Information).

Die-rolling task (Phase 1). The procedure of the studies resembled the one adopted in Study 2 with one main difference in the incentive structure. As in Studies 1a–2, second movers were rewarded (US$0.60) only if their reported outcome matched with that of the first mover. However, the participants were informed that at the end of each session, any money not awarded to the second mover in case he or she did not score a double in the dice rolling would be allocated to an experimental fund to benefit the collective. Dishonest behaviour of second movers thus directly resulted in the depletion of the common good. Specifically, Studies 3 and 4 involved two types of common goods to be exploited by an inflated report of doubles by the second mover. In Study 3, the money in the experimental fund was equally divided and allocated to all participants at the end of the data collection. In Study 4, it was donated to a pro-environmental charity that offsets CO_2 emissions (https://www.cooleffect.org). On average, participants playing the role of first mover in the die-rolling task and subsequently the role of sender in the DG reported that the mission of the charity was moderately important for them (mean = 5.16, s.d. = 1.49), as measured on a seven-point Likert scale (1 = not at all important, 7 = extremely important).

Economic game with a TPP (Phase 2). Afterwards, the participants engaged in a DG with third-party punishment as senders or as the TPP, following the procedure of Study 2. In this phase, we also assessed interpersonal trust towards the unknown receiver (α = 0.93 to 0.95). In addition, we assessed expectations about punishment from the TPP by asking senders to indicate how many lottery tickets they expected the third party to invest to reduce the earnings of the other players in the DG (0–100).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated and analysed during the current studies are publicly available at https://osf.io/fm9b3. Source data are provided with this paper.

Code availability

The code used to analyse the data is publicly available at https://osf.io/p986h.

References

1. Harding, L. What are the Panama Papers? A guide to history's biggest data leak. *Guardian* (5 April 2016).
2. Rose-Ackerman, S. Trust, honesty and corruption: reflection on the state-building process. *Arch. Eur. Sociol.* **42**, 526–570 (2001).
3. Köbis, N. C., Van Prooijen, J. W., Righetti, F. & Van Lange, P. A. M. Prospection in individual and interpersonal corruption dilemmas. *Rev. Gen. Psychol.* **20**, 71–85 (2016).
4. Graycar, A. & Smith, R. G. *Handbook of Global Research and Practice in Corruption* (Edward Elgar, 2011).
5. Mauro, P. Corruption and growth. *Q. J. Econ.* **110**, 681–712 (1995).
6. Rose-Ackerman, S. The economics of corruption. *J. Public Econ.* **4**, 187–203 (1975).
7. Weisel, O. & Shalvi, S. The collaborative roots of corruption. *Proc. Natl Acad. Sci. USA* **112**, 10651–10656 (2015).
8. Gross, J., Leib, M., Offerman, T. & Shalvi, S. Ethical free riding: when honest people find dishonest partners. *Psychol. Sci.* **29**, 1956–1968 (2018).
9. Gründler, K. & Potrafke, N. Corruption and economic growth: new empirical evidence. *Eur. J. Polit. Econ.* **60**, 101810 (2019).

10. Gupta, S., Davoodi, H. & Alonso-Terme, R. Does corruption affect income inequality and poverty? *Econ. Gov.* **3**, 23–45 (2002).
11. Gino, F., Ayal, S. & Ariely, D. Contagion and differentiation in unethical behavior: the effect of one bad apple on the barrel. *Psychol. Sci.* **20**, 393–398 (2009).
12. Köbis, N. C., van Prooijen, J. W., Righetti, F. & Van Lange, P. A. M. The road to bribery and corruption: slippery slope or steep cliff? *Psychol. Sci.* **28**, 297–306 (2017).
13. Rothstein, B. & Eek, D. Political corruption and social trust: an experimental approach. *Ration. Soc.* **21**, 81–112 (2009).
14. Banerjee, R. Corruption, norm violation and decay in social capital. *J. Public Econ.* **137**, 14–27 (2016).
15. Muthukrishna, M., Francois, P., Pourahmadi, S. & Henrich, J. Corrupting cooperation and how anti-corruption strategies may backfire. *Nat. Hum. Behav.* **1**, 0138 (2017).
16. Baumert, A., Halmburger, A., Rothmund, T. & Schemer, C. Everyday dynamics in generalized social and political trust. *J. Res. Pers.* **69**, 44–54 (2017).
17. Balliet, D. & van Lange, P. A. M. Trust, punishment, and cooperation across 18 societies: a meta-analysis. *Perspect. Psychol. Sci.* **8**, 363–379 (2013).
18. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ. Press, 1990).
19. Powers, S. T., van Schaik, C. P. & Lehmann, L. Cooperation in large-scale human societies—what, if anything, makes it unique, and how did it evolve? *Evol. Anthropol.* https://doi.org/10.1002/evan.21909 (2021).
20. Yamagishi, T. The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.* **51**, 110–116 (1986).
21. Hruschka, D. et al. Impartial institutions, pathogen stress and the expanding social network. *Hum. Nat.* **25**, 567–579 (2014).
22. Spadaro, G., Gangl, K., Van Prooijen, J.-W., Van Lange, P. A. M. & Mosso, C. O. Enhancing feelings of security: how institutional trust promotes interpersonal trust. *PLoS ONE* **15**, e0237934 (2020).
23. Macy, M. W. & Sato, Y. Trust, cooperation, and market formation in the U.S. and Japan. *Proc. Natl Acad. Sci. USA* **99**, 7214–7220 (2002).
24. Dinesen, P. T. Where you come from or where you live? Examining the cultural and institutional explanation of generalized trust using migration as a natural experiment. *Eur. Sociol. Rev.* **29**, 114–128 (2013).
25. Sønderskov, K. M. & Dinesen, P. T. Trusting the state, trusting each other? The effect of institutional trust on social trust. *Polit. Behav.* **38**, 179–202 (2016).
26. Gächter, S. & Schulz, J. F. Intrinsic honesty and the prevalence of rule violations across societies. *Nature* **531**, 496–499 (2016).
27. Cohn, A., Maréchal, M. A., Tannenbaum, D. & Zünd, C. L. Civic honesty around the globe. *Science* **365**, 70–73 (2019).
28. Keizer, K., Lindenberg, S. & Steg, L. The spreading of disorder. *Science* **322**, 1681–1685 (2008).
29. Balliet, D. & Van Lange, P. A. M. Trust, conflict, and cooperation: a meta-analysis. *Psychol. Bull.* **139**, 1090–1112 (2013).
30. Fehr, E. & Gächter, S. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000).
31. Molho, C., Tybur, J. M., Van Lange, P. A. M. & Balliet, D. Direct and indirect punishment of norm violations in daily life. *Nat. Commun.* **11**, 3432 (2020).
32. Čábelková, I. & Hanousek, J. The power of negative thinking: corruption, perception and willingness to bribe in Ukraine. *Appl. Econ.* **36**, 383–397 (2004).
33. Bowler, S. & Karp, J. A. Politicians, scandals, and trust in government. *Polit. Behav.* **26**, 271–287 (2004).
34. Halmburger, A., Rothmund, R., Schulte, M. & Baumert, A. Psychological reactions to political scandals: effects on emotions, trust, and the need for punishment. *J. Polit. Psychol.* **2**, 30–51 (2012).
35. Gächter, S., Renner, E. & Sefton, M. The long-run benefits of punishment. *Science* **322**, 1510 (2008).
36. Croson, R. & Gächter, S. The science of experimental economics. *J. Econ. Behav. Organ.* **73**, 122–131 (2010).
37. van Dijk, E. & De Dreu, C. K. W. Experimental games and social decision making. *Annu. Rev. Psychol.* **72**, 415–438 (2021).
38. Soraperra, I. et al. The bad consequences of teamwork. *Econ. Lett.* **160**, 12–15 (2017).
39. Wouda, J., Bijlstra, G., Frankenhuis, W. E. & Wigboldus, D. H. J. The collaborative roots of corruption? A replication of Weisel & Shalvi (2015). *Collabra Psychol.* **3**, 27 (2017).
40. Choshen-Hillel, S., Shaw, A. & Caruso, E. M. Lying to appear honest. *J. Exp. Psychol. Gen.* **149**, 1719–1745 (2020).
41. Gerlach, P., Teodorescu, K. & Hertwig, R. The truth about lies: a meta-analysis on dishonest behavior. *Psychol. Bull.* **145**, 1–44 (2019).
42. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
43. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **25**, 63–87 (2004).
44. Yamagishi, T. & Yamagishi, M. Trust and commitment in the United States and Japan. *Motiv. Emot.* **18**, 129–166 (1994).
45. Capraro, V. & Perc, M. Mathematical foundations of moral preferences. *J. R. Soc. Interface* **18**, 20200880 (2021).
46. Berg, J., Dickhaut, J. & McCabe, K. Trust, reciprocity, and social history. *Games Econ. Behav.* **10**, 122–142 (1995).
47. Van Lange, P. A. M. & Kuhlman, D. M. Social value orientations and impressions of partner's honesty and intelligence: a test of the might versus morality effect. *J. Pers. Soc. Psychol.* **67**, 126–141 (1994).
48. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022); http://www.R-project.org
49. Preacher, K. J. & Hayes, A. F. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. in. *Behav. Res. Methods* **40**, 879–891 (2008).
50. Knack, S. & Keefer, P. Does social capital have an economic payoff? *Q. J. Econ.* **112**, 1251–1288 (1997).
51. Lewicki, R. & Wiethoff, C. in *The Handbook of Conflict Resolution: Theory and Practice* (eds Deutsch, M. & Coleman, P. T.) 104–136 (Jossey-Bass, 2000).
52. Van Lange, P. A. M. Generalized trust: four lessons from genetics and culture. *Curr. Dir. Psychol. Sci.* **24**, 71–76 (2015).
53. Algan, Y. & Cahuc, P. Inherited trust and growth. *Am. Econ. Rev.* **100**, 2060–2092 (2010).
54. La Porta, R., Lopez-de-Silanes, F., Shleifer, A. & Vishny, R. W. Trust in large organizations. *Am. Econ. Rev.* **87**, 333–338 (1997).
55. Van Lange, P. A. M., Joireman, J. & Milinski, M. Climate change: what psychology can offer in terms of insights and solutions. *Curr. Dir. Psychol. Sci.* **27**, 269–274 (2018).
56. van Prooijen, J. W. *The Moral Punishment Instinct* (Oxford Univ. Press, 2017).
57. Stagnaro, M. N., Arechar, A. A. & Rand, D. G. From good institutions to generous citizens: top-down incentives to cooperate promote subsequent prosociality but not norm enforcement. *Cognition* **167**, 212–254 (2017).
58. Marcin, I., Robalo, P. & Tausch, F. Institutional endogeneity and third-party punishment in social dilemmas. *J. Econ. Behav. Organ.* **161**, 243–264 (2019).
59. Kerr, N. L. et al. "How many bad apples does it take to spoil the whole barrel?": social exclusion and toleration for bad apples. *J. Exp. Soc. Psychol.* **45**, 603–613 (2009).
60. Liebrand, W. B. G., Wilke, H. A. M., Vogel, R. & Wolters, F. J. M. Value orientation and conformity: a study using three types of social dilemma games. *J. Confl. Resolut.* **30**, 77–97 (1986).

61. Brohmer, H. et al. Inspired to lend a hand? Attempts to elicit prosocial behavior through goal contagion. *Front. Psychol.* **10**, 545 (2019).
62. Stirrat, M. & Perrett, D. I. Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychol. Sci.* **21**, 349–354 (2010).
63. Erdfelder, E., Faul, F., Buchner, A. & Lang, A. G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).
64. Casler, K., Bickel, L. & Hackett, E. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Comput. Hum. Behav.* **29**, 2156–2160 (2013).
65. Paolacci, G. & Chandler, J. Inside the Turk: understanding Mechanical Turk as a participant pool. *Curr. Dir. Psychol. Sci.* **23**, 184–188 (2014).
66. Hendriks, A. *SoPHIE—Software Platform for Human Interaction Experiments (v3.2.1)* (University of Osnabrück, 2012).

Acknowledgements

We thank D. Balliet, L. Hoeft and the members of the Amsterdam Cooperation Lab for helpful comments on the manuscript. G.S. acknowledges funding from the ERC Starting Grant (no. 635356) awarded to D. Balliet. C.M. acknowledges IAST funding from the French National Research Agency under grant no. ANR-17-EURE-0010 (Investissements d'Avenir programme). A.R. acknowledges funding from the ERC Advanced Grant (no. 785635) awarded to C. K. W. De Dreu. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

G.S., C.M., J.-W.V.P., A.R., C.O.M. and P.A.M.V.L. conceived the project. G.S. collected the data for Studies 1a through 2. C.M. collected the data for Studies 3 and 4. G.S. analysed the data and wrote the initial draft of the manuscript with input and revisions from C.M., J.-W.V.P., A.R., C.O.M. and P.A.M.V.L.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-022-01457-w.

Correspondence and requests for materials should be addressed to Giuliana Spadaro.

Peer review information *Nature Human Behaviour* thanks Valerio Capraro and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

nature portfolio

Corresponding author(s): Giuliana Spadaro

Last updated by author(s): Jun 10, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
□	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
□	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
□	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
□	☒	A description of all covariates tested
☒	□	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
□	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
□	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	□	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	□	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
□	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection: Individual responses were collected using the Qualtrics software (Study 1a: Version May 2017; Study 1b: Version June 2017) and SoPHIE Labs software (Studies 2,3,4: Version 3.2.1).

Data analysis: Data were analyzed using the software R (version 4.0.5). Mediation analyses were conducted using the PROCESS macro (for R version 4.1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The datasets generated and analyzed during the current studies are publicly available at osf.io/fm9b3.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Gender of participants was determined based on self-reporting at the end of each study. Bivariate correlations between gender and study variables are reported in the Supplementary Information. Individual-level gender information is provided in the source data.
Population characteristics	See "Behavioural & social sciences study design" section.
Recruitment	Participants were recruited through the MTurk crowdsourcing online platform. Participation in all studies was compensated through a monetary show-up fee, and the behavior in the games was incentivized through behavior-dependent monetary lotteries conducted at the end of the data collections. To minimize self-selection bias, we did not set specific requirements for participation (besides being located in the United States). In addition, we assigned the survey a generic title unrelated to the research question and hypotheses (i.e., "real-time interaction study on decision-making").
Ethics oversight	All studies were approved by the Scientific and Ethical Review Board (VCWE) of the Faculty of Behavioural & Movement Sciences, VU Amsterdam Application VCWE-2017-085

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	- Study 1a and 1b: Quantitative data, experimental research design; - Study 2 through 4: Quantitative data, correlational research design.
Research sample	All studies involved participants recruited through MTurk. The percentage of women in the sample ranged from 34% to 49%, participants' age ranged from 34.88 (SD = 10.37) to 37.73 (SD = 10.82). The sample was chosen as samples from MTurk are heterogeneous in terms of socio-economic and ethnic diversity, and MTurk is a reliable platform to perform behavioral tasks.
Sampling strategy	Data were collected using a digital convenience sampling strategy through the MTurk crowdsourcing online platform. Sample sizes were pre-determined based on a-priory power analyses performed through G*Power as follows: - Study 1a (experimental): An a-priori power analysis revealed a required sample size of 528 to achieve statistical power of .80 to detect an effect size of d = 0.30 of the corruption manipulation on interpersonal trust (two-sided). - Study 2 (correlational): An a-priori power analysis revealed a required sample size of 380 (190 dyads) to achieve statistical power of .95 to detect an effect size of d = 0.24 (two-sided). - Studies 3 and 4 (correlational): An a priori sensitivity power analysis revealed that a sample size of 100 dyads would give statistical power of .80 to detect an effect size of r = .24 (one-sided). For each study, data collection was stopped once the required target was met.
Data collection	All studies involved online data collection. Participants accessed the link to the experiment on MTurk, and were re-directed to an online study implemented in Qualtrics (Studies 1a through 1b) or in SoPHIE Labs (Studies 2 through 4). In the interactive studies 2 through 4, participants were matched in dyads by the software on a “first-come, first-served” basis, so that each experimental slot was filled by two participants as soon as they provided their informed consent. Thus, across all studies, researchers could not influence the results knowing the hypotheses and/or the experimental conditions in advance.
Timing	Study 1a: May 24th 2017 - May 29th 2017 Study 1b: June 13th 2017 - June 17th 2017 Study 2: October 27th 2017 - November 19th 2017 Study 3a: October 6th 2020 - October 22nd 2020 Study 3b: August 14th 2020 - September 8th 2020
Data exclusions	No individual data were excluded from analyses.
Non-participation	No participants dropped out or declined participation.

Randomization	In Studies 1a and 1b participants were randomly assigned into the between-subjects experimental conditions. Randomization was handled by Qualtrics.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	□ Antibodies
☒	□ Eukaryotic cell lines
☒	□ Palaeontology and archaeology
☒	□ Animals and other organisms
☒	□ Clinical data
☒	□ Dual use research of concern

Methods

n/a	Involved in the study
☒	□ ChIP-seq
☒	□ Flow cytometry
☒	□ MRI-based neuroimaging