Causes and Cures of Corruption

Michael Muthukrishna[1], Patrick Francois[2,3], Shayan Pourahmadi[4] & Joseph Henrich[2,3,5]

[1]London School of Economics and Political Science, London, UK

[2]University of British Columbia, Vancouver, Canada

[3]Canadian Institute for Advanced Research

[4]Columbia University, New York, NY, USA

[5]Harvard University, Cambridge, MA, USA

Understanding how humans sustain cooperation in large, anonymous societies remains a central question of theoretical and practical importance. Building on foundational work using peer punishment in the Public Goods Game, research suggests that cooperation among strangers can be sustained by institutional punishment, whereby costly punishment is meted out by a designated leader using pooled resources[1,2]. Such institutional punishment is analogous to governments, police forces, and other institutions that sanction free-riders on behalf of individuals in large societies. In the real world, however, corruption can undermine the effectiveness of these institutions with accompanying economic and social costs[3-5]. Greece, for example, has the highest level of corruption in the European Union, with estimates placing it near China and Brazil[6]. Levels of corruption correlate with institutional, economic, and cultural factors, but the causal direction of these relationships is difficult to determine[5,7]. Here we theoretically and experimentally model corruption and cooperation, testing the effect of a leader's power to punish, economic potential, transparency, leader investment in the public good, and cultural background. Consistent with theoretical predictions, our findings reveal four patterns: (1) corruption causes a large (25%) decrease in public good provisioning; (2) empowering leaders decreases cooperative contributions when bribery is an option—in contrast

to typical Public Goods Game results, where stronger leaders increase contributions; (3) growing up in a more corrupt society led to more acceptance of bribes in our experiments, suggesting that people internalize norms of corruption; and, (4) anti-corruption strategies, such as transparency and forcing leaders to invest in the public good are effective under some conditions, but can increase corruption when leaders are weak and economic potential is poor. These results suggest that a more nuanced approach to corruption is needed, and that proposed panaceas, such as transparency, may be harmful in some contexts.

Cooperation, particularly large-scale anonymous cooperation, remains an important puzzle to both evolutionary and social scientists, with real world social and economic implications. One method for sustaining cooperation that has received considerable attention is costly punishment[8-10], whereby individuals pay a cost to punish free riders who fail to contribute to the public good. While cross-cultural evidence shows the ubiquity of costly punishment in large-scale societies (though not in small-scale societies), there is some variability in both the motivation to punish free riders and the willingness to punish cooperators[11-13].

Research on the role of peer punishment in sustaining cooperation reveals two major challenges: (1) the second order free-rider problem in which individuals defect on the job of punishing and thereby increase their payoffs[14,15], and (2) the problem of counter-punishment—punishment as revenge for previously being punished[9,16]. Institutional, or pool, punishment resolves these problems by designating one individual as a "Leader" who can extract taxes and punish free-riders on behalf of other players. Institutional punishment reduces the problems of both second-order free-riding and counter-punishment, and may thus be important in explaining the emergence and sustainment of large-scale cooperation[2]. Moreover, recent empirical research shows that participants (at least WEIRD participants[17]) prefer institutional punishment to peer punishment[1,18].

Institutional punishments, as typically modelled in PGGs, serve to incentivize player choices when contributing to the public pool, and work by constraining leader choices to either punishing players or doing nothing. In the real world, however, channels such as bribery, nepotism, and lobbying allow individuals (or corporations) to avoid contributing to the public pool (e.g., evading taxes) as well as avoid being punished (e.g., by paying a bribe instead)—real leaders and institutions are corruptible.

To model corruption, we modified the repeated institutional punishment PGG (IPGG from herein), giving players an additional choice in which they can improve the leader's payoff using their own resources (i.e., offer a bribe, though we use neutral language). In turn, leaders have an additional choice for each player: leaders can choose to take the contribution (i.e., accept the bribe) or not. In this Bribery Game (BG), we manipulate the pool multiplier (a proxy for economic potential) and the punishment multiplier (power of the leader to punish). We also introduce three corruption mitigation strategies: partial transparency (revealing leader contributions), full transparency (revealing all leader behaviour, including bride taking), and leader investment (forcing leaders to contribute their endowment to the public pool).

We analyse the evolutionary dynamics of the IPGG and BG to better understand player behaviour under different treatments (see SI for details). We model an IPGG with a fixed tax rate to more realistically model real world institutions, where taxes are not directly correlated with punishment and where leaders punish without a large cost to themselves (since their own taxes are a small part of the taxes contributing to the pool punishment or institution). We then modify this game to a BG by offering players and leaders the choice to offer and accept bribes. It is easy to see that without punishment, contributions will tend toward zero—contribution levels are contingent on the strength of leaders and their tendency to punish low contributors. Our theory predicts that leaders will use taxes to punish, since these are not personally costly and benefit the leader's payoff by increasing the size of the public good. Increased leader strength predicts higher contributions and more public good provisioning. Our analysis of the BG modification demonstrates that players have no incentive to offer contributions or bribes unless they are punished for not doing so. However, when bribery is an option, leaders have a greater incentive to punish people for not offering brides than for not contributing. More power gives leaders an increased ability to impose their will, increasing the rate of bribery at the expense of the public good. Thus, in contrast to the IPGG, stronger leaders in the BG should *reduce* contributions and public good provisioning. However, if players have a preference for contributions over bribes (such as when potential returns on the public good are higher or anti-corruption norms exist), then the incentive to punish brides over contributions is dampened. For full set of predictions, see SI.

Cost of corruption. We find that mean contributions drop by 25% when bribery is an option (BG compared to IPGG). The difference between these conditions represents a 0.43 [CI_{95}: -0.49,-0.38] standard deviation loss (1.4 points per period; equivalent to 14% of initial endowment or $2.10 over the course of the game). Not surprisingly, when corruption can enter, it does, and cooperation declines.

Causes of corruption. Leaders with a stronger punishment multiplier at their disposal were about twice as likely to accept bribes and about 3 times less likely to do nothing (Table 1). In contrast, when accepting bribes is not an option (IPGG), these more powerful leaders were as likely or slightly more likely to do nothing (see SI). As predicted, more power led to more corrupt behaviour.

Norms affect behaviour: more bribe acceptance was observed when players grew up in more corrupt countries. For every one standard deviation increase in players' exposure corruption scores, leaders were 1.2 times more likely to accept a bribe. In contrast, when players' parental heritage included countries with higher corruption norms, leaders were 1.5 times *less* likely to accept bribes for every standard deviation increase in corruption score, and 1.6 times more likely to do nothing (see Table 1; SI shows all models). Our results offer some evidence that norms of corruption are internalized and affect economic behaviour.

	Accept Bribe	**Punish**	**Do Nothing**
High Economic Potential	1.37 [0.65,2.21]	0.79 [0.41,1.14]	0.81 [0.29,1.40]
Strong Leader	2.14 [1.18,3.36]	1.08 [0.60,1.61]	0.29 [0.10,0.50]
Player Exposure Corruption Score	1.22 [1.01,1.44]	0.99 [0.81,1.19]	0.79 [0.63,1.02]
Player Heritage Corruption Score	0.65 [0.54,0.79]	1.17 [0.96,1.40]	1.55 [1.25,1.89]
(Intercept)	0.57 [0.05,1.54]	0.16 [0.02,0.39]	3.01 [0.12,9.50]
Obs.	1396	1396	1396
N	175	175	175
Groups	45	45	45
DIC	36.13	18.23	18.45

Table 1 | Leader Decision Each column reports the odds and 95% confidence interval of the behaviour in the column heading (e.g. Accept Bribe in column 1) compared to engaging in one of the other two behaviours (e.g. Punish or Do Nothing for column 1). The odds are estimated using an MCMC categorical GLMM, with the behaviour coded as 1 and the other two behaviours coded as 0. The confidence intervals are Highest Posterior Density (HPD) intervals. Each model regresses the behaviour in the Bribery Game (with no transparency or leader investment) on economic potential (low vs. high), leadership strength (weak vs. strong), and both player's and leader's exposure corruption score (z-score) and heritage corruption score (z-score), controlling for period, order of conditions, order of background questions, group size, age, and gender with random effects for individuals within groups. Here we report only the predictors of interest. The full model is reported in the SI.

Cures for corruption. Having generated corruption, we attempted to suppress it by modifying the BG using two different forms of transparency measures and by forcing leaders to invest in the public good. The first transparency approach, Partial Transparency, allowed all players to see the leader's contribution, thereby offering leaders an opportunity to establish or reveal a norm by revealing to players how much or how little leaders invested in the public pool. The second transparency approach, Full Transparency, allowed players to see all leader actions: leader contributions, the anonymized contributions and bribes from each player, and the leader's decision in each case. Finally, Leader Investment forced leaders to maximally contribute their endowment to the public good, thereby tying a large part of their payoff to the efficiency of the public good and incentivizing them to increase other's contributions rather than accept bribes. Tying leader payoffs to the success of the public good has explicitly been used as an anti-corruption measure in places such as Singapore, which has one of the lowest levels of corruption (based on Corruption Perception Index[6]) and the highest paid leader in the world[19]. Singaporean minister salaries are pegged at the salaries of top professionals and Singapore's Gross Domestic Product.

To determine the effectiveness of these anti-corruption measures, we compared contributions in each condition to the IPGG (control) and to the BG. We regressed contributions (z-scores) on treatment, economic potential, and leader strength. Figure 1 summarizes the results of this regression and reports separate coefficients within each condition. Note that these values come from a single model and are calculated by changing reference groups (see SI). Raw mean contribution values are graphed in Figure 2.

	Weak Leaders		Strong Leaders	
Poor Economic Potential	Control	BG	Control	BG
Control	-	0.21**	-	0.52***
Bribery Game (BG)	-0.21***	-	-0.53***	-
BG + Partial Transparency	-0.31***	-0.10*	-0.53***	-0.01
BG + Full Transparency	-0.20***	-0.01	-0.06	0.47***
BG + Leader Investment	-0.46***	-0.25***	-0.17**	0.36***
Rich Economic Potential	Control	BG	Control	BG
Control	-	0.39***	-	0.57***
Bribery Game (BG)	-0.39***	-	-0.57***	-
BG + Partial Transparency	-0.30***	0.09^{+}	-0.44***	0.13**
BG + Full Transparency	-0.15**	0.24***	-0.25***	0.32***
BG + Leader Investment	-0.15**	0.24***	-0.21***	0.36***

***** p < 0.001** **** p < 0.01** *** p < 0.05** **+ p < 0.10**

Figure 1 | Cures for corruption Coefficients and colours indicate the effect on contributions, on public goods provisioning. Deeper blue shading indicates greater public goods provision and darker red indicates reduced public goods. All coefficients are extracted from a single model by changing reference groups. The 4 large rectangles show the effect of each treatment with weak and strong leaders (columns) and poor and rich economic potential (rows). Within each rectangle, the columns represent the reference group treatment (Control, BG). Each row reports the coefficient of each treatment compared to this reference group. Contributions are z-scores, so coefficients represent standardized differences. The full model is reported in the SI. Note that in all treatments and structural contexts, the BG has lower contributions than the structurally equivalent IPGG (control). Corruption mitigation effectively increases contributions (though not to control levels) when leaders are strong or economic potential is rich. When leaders are weak and economic potential is poor, the apparent corruption mitigation strategy Full Transparency has no effect and Partial Transparency and Leader Investment further *decrease* contributions.

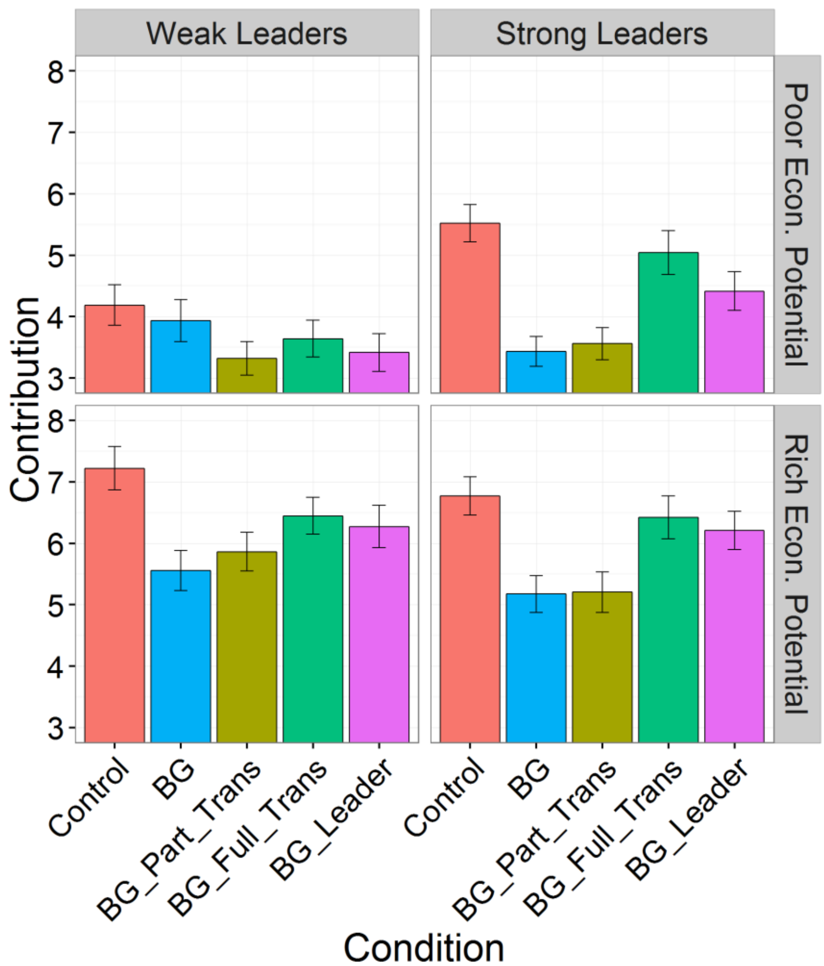

Figure 2 | Contributions by condition Raw contributions for each within-subjects treatments in between-subjects structural contexts with 95% confidence intervals. These data are consistent with our theory, which predicts that more powerful leaders increase contributions in the IPGG, but decrease contributions in the BG.

Figures 1 and 2 reveal that stronger leaders are better able to increase the efficiency of public goods provisioning when economic potential is poor and bribery is not an option (red bars in top row), but when bribery is an option (blue bars), stronger leaders in poor contexts reduce the efficiency of the public good, making themselves wealthy at the expense of other players. However, Partial Transparency, Full Transparency, and Leader Investment all increased contributions, here and in all contexts, except when leaders were weak and economic potential was poor. When leaders were weak and economic potential poor, Full Transparency had no effect on contributions. In these conditions, Partial Transparency and Leader Investment *decreased* contributions to levels lower than the standard BG—that's they led to less public good provisioning.

Although the cost of bribery was seen in all contexts, in poor economic contexts, the already low contributions were reduced even further. That is, even if powerful leaders are accepting bribes at comparable levels in both poor and rich economic contexts, the degree of corruption may not be as visible if economic potential is high. Leaders in richer economic contexts, like the United States, may accept "bribes" in the form of lobbying or campaign funding, which may indeed reduce the efficiency of the public good, but this cost isn't as obvious since economic potential is already much higher than in other nations. In contrast, in the Democratic Republic of Congo, a poorer economic context, corruption further reduces the already low public good provisioning. Unfortunately, our results suggest that in these contexts with weak institutions and poor economic potential, efforts to mitigate corruption, such as transparency or leader investment, could backfire, further *reducing* investments in the public good. These results reflect leaders lacking the power to increase contributions through punishment and thus recouping the cost of their investment in the public good by accepting bribes and lacking this power, revealing a low contribution norm. Thus the lessons in fighting corruption when institutions have the power to sustain public goods (if only corruption were reduced) and the potential for economic growth is high, may not only fail to apply when these conditions are not met, but can worsen the situation.

The increase in corrupt behaviour with direct exposure to corrupt norms is consistent with the internalization of perceived norms[20,21] and with previous empirical data showing, for example, diplomats from high-corruption countries accumulating more unpaid parking violations[22]. However, the decreased probability of accepting bribes among those whose cultural background includes more corrupt countries, suggests that second-generation and later migrants are not as corrupt as their peers from less corrupt nations. This may represent the self-selection of immigrants from their home countries or may be a form of "identity denial"[21], whereby acculturated individuals actively avoid the stereotypes of their inherited culture. Although we have a large range of corruption scores (see SI), our sample is limited to migrants in a Canadian context and further investigation is required to determine if these cultural results generalize. Together these results suggest that corruption may be rooted in structural factors, but that internalized corruption norms may cause these behaviours to persist.

Overall, these results suggest that: (1) stronger institutions and leaders are required to sustain public goods contributions when economic potential is poor and the incentive to free-ride is high; (2) in this context, when able to, leaders will abuse their power with a noticeable economic cost to corruption. However, (3) even if economic potential is poor, if leaders are powerful, anti-corruption measures can be effective at increasing public good provisioning. Thus, efforts to mitigate corruption in poorer economic contexts must go hand-in-hand with strengthening institutions. When leaders have less punitive power, efforts such as transparency may have no effect or even decrease contributions since they reveal the rationality of low public good contributions and that most leaders do not contribute.

Though these experimental results begin to offer insights into the causal effect of corruption on cooperation, extending such experimental findings demands great caution. Laboratory work on the causes and cures of corruption must inform and be informed by real world investigations of corruption from around the globe. Thus, aiming only to drive future investigation, our results suggest that as economic potential grows, less government intervention is required to enforce cooperation and increased power may be misused, requiring greater anti-corruption efforts. In contrast, when economic potential is poor, strong government intervention is most effective at decreasing free-riding, as long as this intervention is paired with strategies to mitigate corruption. This may help explain why intuitions about "cures for corruption" based on experiences in rich nations do not work as well in poorer nations.

Methods

Recruitment of Subjects. We had a total of 273 participants (166 female, mean age 20.91) drawn from an Economic Subject Pool open to the general public. Participant ethnic backgrounds were as follows: 63 Euro Canadians, 158 East Asians, 16 South Asians, 36 Other Ethnicities. Participants played in groups of between 4 and 7 players.

Experimental Design. We used a 2 (high vs low economic potential) x 2 (weak vs strong leader power) between-subjects experimental design with 5 within-subjects treatments (institutional punishment public goods game; control, bribery game, bribery game with partial transparency, bribery game with full transparency, bribery game with leader investment. In the real world, leaders make institutional decisions based on a fixed budget to which they are one among many contributors and which has to be spent. To better model these conditions, we extracted fixed taxes for punishment, which were discarded if not used. Participants were randomly assigned to one of the 4 between-subjects treatments and 4 of the 5 within-subjects treatments.

To test possible contributing causes of corruption, we randomly assigned each group of participants to a treatment with either high or low (1) marginal per capita rate of return (0.3 vs. 0.6) as a measure of economic potential and (2) a punishment multiplier (1 vs. 3) as a measure of the strength of the leader or institution. The marginal per capita rate of return is the expected return for every point invested in the public pool and the punishment multiplier is the number of points subtracted from a sanctioned player for every tax point spent on punishing that player.

Within-subjects treatments were played in a random order with pre-recorded video instructions prior to each period. A quiz was given prior to beginning, to ensure participants knew how each treatment worked. We used a block randomization design, where participants played a minimum of 10 rounds, but the game may have ended at any point prior to the completion of 10 rounds. At 10 rounds, participants were informed which round the period ended or played further rounds until the game ended. In this way, we had 10 rounds to analyse without end game effects—participants did not know when the game would end. Participants were paid for 10 random rounds from across all conditions. They were paid at a rate of 15c per point, with a show up fee of $10.

Measures. We measured Age, Gender, University Degree or Occupation and Major or Industry, Prestige/Dominance, Right Wing Authoritarianism, whether they spent their entire life in Canada, where else they've lived, what suburb they grew up in, ethnic group, Religion and importance of religion, how well they speak their native language (Cultural Competence), Inclusion of Other in the Self Scale (Identification with their Ethnic Group and Identification with Canadians), Vancouver Index of Acculturation, and Mainstream vs Heritage Acculturation (Integration into Culture). Two corruption scores were calculated for each person using the mean perception of corruption index from Transparency International for all of the countries the participant have lived in and all countries from which they derived their ethnic heritage. The corruption index begins at 0 (most corrupt) up to 100 (least corrupt). For each country, we subtracted this value from 100 (so that higher scores indicated higher corruption). Perception of corruption was chosen as the measure of corruption as it indicated the perceived norm for national corruption.

The heritage corruption score primarily represents the potential influence of vertically transmitted corruption norms (parent to child), whereas the exposure corruption score represents corruption norms which the participant may have acquire through non-parental cultural transmission or direct experience.

We asked the last 39 groups (194 participants) their preferences for the conditions of the game. These participants were asked these questions after taking all other measures so that they were no different to the preceding 17 groups (79 participants). We report these preferences, along with details of all measures in the SI.

References

1 Fehr, E. & Williams, T. *Endogenous emergence of institutions to sustain cooperation* (2013).
2 Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466**, 861-863 (2010).
3 Ambraseys, N. & Bilham, R. Corruption kills. *Nature* **469**, 153-155 (2011).
4 Kaufmann, D. Myths and realities of governance and corruption. *Available at SSRN 829244* (2005).
5 Treisman, D. What have we learned about the causes of corruption from ten years of cross-national empirical research? *Annu. Rev. Polit. Sci.* **10**, 211-244 (2007).
6 Transparency International. (2014).
7 Treisman, D. The causes of corruption: a cross-national study. *Journal of public economics* **76**, 399-457 (2000).
8 Fehr, E. & Gachter, S. Cooperation and Punishment in Public Goods Experiments. *American Economic Review* **90**, 980-994 (2000).
9 Gächter, S., Renner, E. & Sefton, M. The long-run benefits of punishment. *Science* **322**, 1510-1510 (2008).
10 Rockenbach, B. & Milinski, M. The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718-723 (2006).
11 Henrich, J. *et al.* Costly punishment across human societies. *Science* **312**, 1767-1770 (2006).
12 Henrich, J. *et al.* Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480-1484 (2010).
13 Herrmann, B., Thöni, C. & Gächter, S. Antisocial Punishment Across Societies. *Science* **319**, 1362-1367, doi:10.1126/science.1153808 (2008).
14 Boyd, R. & Richerson, P. J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and sociobiology* **13**, 171-195 (1992).
15 Panchanathan, K. & Boyd, R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* **432**, 499-502 (2004).
16 Nikiforakis, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* **92**, 91-112, doi:http://dx.doi.org/10.1016/j.jpubeco.2007.04.008 (2008).
17 Henrich, J., Heine, S. J. & Norenzayan, A. Most people are not WEIRD. *Nature* **466**, 29-29 (2010).
18 Traulsen, A., Röhl, T. & Milinski, M. An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proceedings of the Royal Society B: Biological Sciences*, rspb20120937 (2012).
19 Kuan Yew, L. *From Third World to First World; the Singapore Story: 1965–2000*. (Harper Collins Publishers, 2000).
20 Chudek, M. & Henrich, J. Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences* **15**, 218-226 (2011).
21 Chudek, M., Muthukrishna, M. & Henrich, J. in *The Handbook of Evolutionary Psychology* Vol. 2 (ed David M. Buss) Ch. 30, (John Wiley and Sons, 2015).

22 Fisman, R. & Miguel, E. Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political economy* **115**, 1020-1048 (2007).

Supplementary Information is available in the online version of the paper. Data and analytic code is available at DataDryad (ADD LINK).

Acknowledgements J.H. and P.F. acknowledges support from the Canadian Institute for Advanced Research.

Author Contributions M.M., P.F., S.P., and J.H. developed theory, designed the experiments, and wrote the paper. M.M. and S.P. carried out the experiments. M.M., S.P., and J.H. conducted the statistical analyses.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.M. (m.muthukrishna@lse.ac.uk).