



## U4 BRIEF 2023:1

Caryn Peiffer  
Nic Cheeseman  
Series editor  
Peter J. Evans

# Message misunderstood: Why raising awareness of corruption can backfire

---

**Corruption erodes sustainable and inclusive development. It is both a political and technical challenge. The U4 Anti-Corruption Resource Centre (U4) works to understand and counter corruption worldwide.**

**U4 is part of the Chr. Michelsen Institute (CMI), an independent development research institute in Norway.**

[www.u4.no](http://www.u4.no)

[u4@cmi.no](mailto:u4@cmi.no)

---

## **Read online**

[www.u4.no/r/BR2301](http://www.u4.no/r/BR2301)

**Anti-corruption strategies almost always contain an awareness raising element. A growing body of research has examined what impact such anti-corruption messaging is having. Findings suggest that resources spent on raising awareness with anti-corruption messaging may be wasted, or even do more harm than good. Research is yet to identify a communications strategy that will always work as intended. It is essential that all anti-corruption messaging learns lessons from other fields, and is always tested before being rolled out.**

before they are communicated to the wider public.

- Pre-deployment testing must be experimental. Only experimental tests provide a systematic estimate of the effect of exposure to messages.
- This kind of pre-deployment message testing should be standard when developing anti-corruption awareness raising campaigns.

## **Main points**

- A growing body of research has examined the effectiveness of anti-corruption awareness raising messaging.
- Research has shown that messaging focused on the scale and consequences of corruption is especially prone to 'backfire' by reinforcing beliefs that the problem is too big to solve.
- Even 'upbeat' messaging emphasising progress made in controlling corruption has been found to make the situation worse, suggesting that a range of anti-corruption awareness raising communication may do more harm than good.
- Many studies have also found that a variety of messages have no impact, suggesting that awareness raising efforts risk wasting resources.
- The only way to make sure that anti-corruption communication campaigns avoid wasting money or backfiring is to test their impact on a sample of the target audience

# Contents

<b>Awareness raising as an anti-corruption intervention</b>	<b>5</b>
<b>Potential to backfire</b>	<b>6</b>
<b>No safe bets</b>	<b>11</b>
<b>The need to test messages</b>	<b>14</b>
<b>References</b>	<b>16</b>

# Awareness raising as an anti-corruption intervention

Anti-corruption strategies almost always contain an awareness raising element.<sup>1</sup> Governments, donors and/or non-governmental organisations routinely aim to raise public awareness about the problem and consequences of corruption through billboards, posters, murals, and radio and television shows. Along with hoping that messaging will increase reports of and refusals to engage in corruption, many also see awareness raising as a tool to shift social norms and tolerance towards corruption, as well as to strengthen grassroots demands for accountability and clean government.

In practice, the effectiveness of awareness raising efforts is usually assessed in terms of their *reach* – how many people see them – which speaks to an overarching assumption: these campaigns will work as intended as long as enough people are exposed to them. Over the last decade, however, a growing body of research has explored the potential positive and negative impacts of anti-corruption messaging by examining how exposure to messages shapes attitudes and behaviour.

## Messaging can backfire and even make individuals more likely to pay a bribe.

Findings suggest that in some cases such messages *backfire*, even making individuals more likely to pay a bribe, and that when this does not happen, they tend to have no impact at all.<sup>2</sup> Significantly, these kinds of worrying outcomes are not limited to anti-corruption messaging. A wider body of literature on public health campaigns has also highlighted the potential for awareness raising messages to be ineffective or to backfire.<sup>3</sup> Taken together, these alarming findings suggest that resources spent on awareness raising risk being wasted, or even worse, doing more harm than good. Therefore, a clear policy implication has emerged: awareness raising campaigns, including traditional anti-corruption messaging, should be halted until they have been systematically tested and shown to likely have a positive effect.

---

1. Cheeseman and Peiffer 2021; 2022; UNCAC 2004.

2. This brief's literature review draws on Cheeseman and Peiffer (forthcoming).

3. Greszczuk 2020; Steadetal 2019.

# Potential to backfire

Many studies have been conducted on the impact of anti-corruption messages, but prior to 2016 these tended to be more descriptive and did not systematically investigate the impact of messages, or they focused on the impact such messaging had on voting. Since 2016, however, nine prominent studies have been conducted that use experimental techniques to isolate what impact anti-corruption messaging has on beliefs about corruption and anti-corruption efforts, and on willingness to bribe.<sup>4</sup> Each have adopted a similar approach. Participants are randomly assigned to either treatment or control groups, with the former exposed to a message, while the latter is not. Estimates of the impact of exposure to messaging are established by comparing how participants who are exposed to a message (treatment group) respond to a survey or behave in a simulated 'bribery game' to the behaviour or responses of participants who are not shown a message (control group). Table 1 below shows that across the nine studies, 19 messages have been tested in seven countries.

Strikingly, the literature finds that more than half (10) of the 19 messages have backfired to some extent. Specifically, exposure to the message tested in Corbacho et al.'s study in Costa Rica – about bribery increasing – elicited greater self-reported willingness to bribe.<sup>5</sup> Peiffer's four very different messages tested in Jakarta, Indonesia, all increased concern about corruption, reduced pride in the government's anti-corruption response, reduced confidence that ordinary people could easily engage in civic anti-corruption activity, and reduced willingness to protest against corruption.<sup>6</sup> These findings are noteworthy because the four messages tested were different from each other but tended to have the same level of impact. Two of these messages described the widespread nature of corruption, but another touted the government's successes in controlling corruption and the last was even an 'up-beat' message which described ways citizens could get involved in anti-corruption activity.

Finally, and perhaps most worryingly, Cheeseman and Peiffer's study in Lagos, Nigeria, found that messages could make individuals more willing to pay a bribe.<sup>7</sup> What makes this study distinctive is that participants were invited to play a game with real money in which participants had the option of bribing a third-party referee

---

4. Corbacho et al. 2016; Peiffer 2017; 2018; Peiffer and Walton 2022; Köbis et al. 2019; Cheeseman and Peiffer 2020; 2021; 2022; 2022b; Agerberg 2021.

5. 2016.

6. 2017; 2018.

7. 2021.

in order to do better than a rival player. Their findings make for stark reading: three of the five messages tested were found to actually encourage bribery.

This is particularly significant because the messages tested by Cheeseman and Peiffer were also substantively different to one another, demonstrating again that a wide range of messages can have a negative effect. The three that encouraged bribery respectively described corruption as endemic, as being against religious moral teachings, and underlined the government's successes in controlling corruption. Exposure to four of their five messages also reduced agreement that taxes should be paid.<sup>8</sup>

These findings are a stern reminder that raising awareness of corruption through messaging is not the same as helping to address it, and that such efforts may in fact be exacerbating the problem. So why does this happen? Drawing on insights from political messaging and social norms research, scholars have suggested that anti-corruption messaging may be ill equipped to change how people think about corruption because it tends to be the type of issue that people already have strong feelings about. Especially for those already convinced that corruption is widespread, instead of changing minds, messaging is more likely to reinforce pre-existing beliefs that the problem is too big to solve and too intractable to try to resist.<sup>9</sup>

This is especially said to apply to messages which emphasise a *descriptive social norm*; i.e. norms based on beliefs about how others behave.<sup>10</sup> This is worrying, as these are the most common kinds of messages used in practice. A central theme of anti-corruption messaging campaigns has been to raise public awareness to the scale and consequences of corruption.<sup>11</sup> Such messaging invariably highlights a descriptive norm that many people in society engage in, benefit from, and/or facilitate corruption.<sup>12</sup>

### **‘Descriptive norm’ awareness raising risks making corruption appear to be socially acceptable.**

In addition to making people think more about the problem, ‘descriptive norm’ awareness raising (emphasising how widely practiced corruption is) also risks making corruption appear to be socially acceptable. As a result, it can encourage individuals – who have been shown in a number of studies to opt into forms of

---

8. Cheeseman and Peiffer 2022.

9. Peiffer 2018; Cheeseman and Peiffer 2022.

10. eg Legros & Cislighi, 2020.

11. eg UNCAC 2004, Article 13, p. 15.

12. Agerberg 2022.

behaviour that they believe their peers are already doing<sup>13</sup> – to adopt more positive attitudes towards participating in graft.

Worryingly, a similar risk has been noted with respect to raising awareness to other types of social bads. For example, Paluck and Ball<sup>14</sup> argue that descriptive norm-based awareness raising around rape in the Democratic Republic of Congo may have made gender-based violence seem more socially normative, working to encourage men to commit such violence. Six of the 19 messages tested in the anti-corruption messaging literature overtly emphasise a ‘negative’ descriptive norm, that corruption is widely practiced or on the rise, most of which were found to have backfired to some extent (Table 1).

---

13. Tankard and Paluck 2016; Biccheri and Dimant 2019.

14. 2010.



**Table 1: Anti-corruption messaging literature finds most messages counterproductive (red) or ineffective (amber). Few worked as intended (green)**

Themes of messages tested	Effect	Location	Study
Increasing rate of bribery in country	Red	Costa Rica	Corbacho et al. (2016)
Grand corruption is endemic	Red	Jakarta	Peiffer (2017; 2018)
Petty corruption is endemic	Red		
Government successes in anti-corruption	Red		
Citizens can get involved in anti-corruption	Red		
Corruption is endemic	Amber	Port Moresby	Peiffer and Walton (2022)
Corruption is illegal	Amber		
Corruption is against religious teachings	Amber		
Corruption is a 'local' issue	Green		
Bribery declined recently in region	Amber	Manguzi	Köbis et al. (2019)
Corruption is endemic	Red	Lagos	Cheeseman and Peiffer (2021; 2022)
Government successes in anti-corruption	Red		
Corruption is against religious teachings	Red		
Corruption steals tax money	Red		
Corruption is a 'local' issue	Red		
Citizens strongly condemn corruption	Green	Mexico	Agerberg (2021)

Themes of messages tested	Effect	Location	Study
Corruption is endemic	Amber	Albania	Cheeseman and Peiffer (2022b)
Citizens strongly condemn corruption	Amber		
Wealth is lost to other countries	Amber		

**Red:** at least one unwanted effect; **Amber:** no impact/largely no impact across outcomes; **Green:** clear intended impact

\* Specifically, the message tested had no impact bribery in game for those who took on role of citizen. Reduced bribery for those who took on role of public official.

\*\* Specifically, these messages had no impact on self-reported willingness to bribe or report corruption, or on beliefs about acceptability of corruption, though increased agreement that respondents can hold public officials accountable by not voting for corrupt ones, as well as willingness to vote for candidates who focused their campaign on fighting corruption.

## No safe bets

One clear lesson has emerged from the research conducted so far: anti-corruption awareness raising campaigns should avoid emphasising a descriptive norm, as such messaging is especially poised to backfire. Unfortunately, a complementary lesson pointing to messaging strategies that are a safe bet and will likely work as intended has not emerged. Three trends in the research are worth noting in this regard.

First, other types of messages – which do not explicitly focus attention on the scale of the problem – have also been found to backfire. Perhaps most surprisingly, Peiffer<sup>15</sup> and Cheeseman and Peiffer's<sup>16</sup> 'good news' messages emphasising government successes in controlling corruption, and Peiffer's<sup>17</sup> 'upbeat' message about how easily citizens can get involved in anti-corruption civic activity were also found to have the opposite impact to that which was intended. Indeed, these findings are consistent with wider research on so-called 'boomerang effects' of messaging, which has shown that even carefully designed messaging can inadvertently focus attention on information which causes an individual to respond in an unintended way.<sup>18</sup> Importantly for practitioners, these findings suggest that even positive messages can encourage people to recall firmly held beliefs that corruption is endemic and intractable, and so still risk doing more harm than good.

### Messages that lack everyday validity can have unintended effects.

It is also important to recognise that in many contexts 'positively' framed anti-corruption messages will lack credibility. Efforts to raise awareness of the fact that the government is actively controlling corruption, for instance, will lack credibility in a place where no serious efforts have taken place and citizens experience graft on a daily basis. Intuitively, political messaging research has established that messages with weak arguments – and messages that lack everyday validity – are more likely to result in unintended, 'boomerang' effects.<sup>19</sup>

The second trend is that *many messages have failed to register any impact on the outcomes scrutinised*. Peiffer and Walton, for example, found that exposure to three of the four messages tested in their study in Port Moresby, Papua New Guinea – one

---

15. 2017; 2018.

16. 2021; 2022.

17. 2017; 2018.

18. Bryne and Hart 2016.

19. Park et al. 2007.

on corruption as endemic, another about corruption being illegal, and a third framing corruption as against religious, moral teachings – had no impact on attitudes towards reporting corruption.<sup>20</sup> Similarly, Köbis et al. found that exposure to even an up-beat message about how bribery had fallen locally, failed to have any impact on the participants' willingness to pay a bribe in a lab-based simulated bribery game, when those participants took on the role of the 'citizen'.<sup>21</sup> Cheeseman and Peiffer also found that exposure to a message about corruption being a 'local issue' and another which framed corruption as stealing tax money in Lagos had no impact on the majority of participants' willingness to bribe in their bribery game.<sup>22</sup>

Finally, Cheeseman and Peiffer's study in Albania finds that exposure to two messages, one of which described corruption as being endemic, and another with a more positive spin – namely that citizens strongly disapprove of corruption – had no impact on self-reported willingness to bribe or report corruption, or on beliefs about the acceptability of corruption.<sup>23</sup>

In other words, a wide range of messages with different thematic emphases and tones have been found to be unimpactful. These 'null' results are important because the outcomes scrutinised in the literature reflect some of the core aims of anti-corruption awareness raising, such as bolstering positive attitudes about and a willingness to report corruption, reducing the acceptability of corruption, as well as reducing a willingness to engage in bribery. These findings suggest that, in practice, investments in anti-corruption strategic messaging are often a waste of resources.

Third, it seems unlikely that the findings of the two studies that find that messages have worked as intended can be reliably generalised. To be clear, only two (of 19) messages tested thus far have registered only 'positive' or intended impacts. In both cases, subsequent studies testing similar messages have failed to support these positive results, undermining confidence in how generalisable initial optimistic findings were. In the first instance, Peiffer and Walton's previously mentioned study found that exposure to a message framing anti-corruption as 'local' issues encouraged favourable attitudes towards reporting. However, Cheeseman and Peiffer's study tested the impact of a very similar message in Lagos where ethnic identification is also salient, and, as noted earlier, found that it had no impact on bribery.<sup>24</sup>

---

20. 2022.

21. 2019.

22. 2021.

23. 2022b.

24. 2020.

The second message tested which has had a clear ‘positive’ impact was from Agerberg’s survey experiment conducted in Mexico.<sup>25</sup> This message emphasised the fact that citizens strongly condemn corruption. Those exposed to this message demonstrated less acceptance that corruption is a basic part of Mexican culture, and a lower likelihood of self-reported willingness to pay a bribe. However, Cheeseman and Peiffer also tested the impact of a similar anti-corruption message in their survey experiment in Albania.<sup>26</sup> As noted earlier, they found that exposure to it had no impact on self-reported willingness to bribe or report corruption, or on beliefs about acceptability of corruption, suggesting that Agerberg’s hopeful findings may not be generalisable. The overwhelming takeaway from existing research, therefore, is that anti-corruption messages rarely, if ever, have the desired effect.

More broadly with respect to generalisability, it is important to note that awareness raising efforts vary considerably in practice and that the research conducted so far has not considered the impact of some of these variations. For example, it is unclear how the impact of messaging is shaped by who delivers the message, the length and intensity of exposure to messaging, the media through which it is communicated, and how messaging campaigns interact with other anti-corruption efforts. The geographical coverage of studies is also not exhaustive. These limitations point to the need for further research in this area. We currently have no evidence to suggest, however, that these variations matter for the efficacy of messaging campaigns. The need for further research therefore does not take away from the clear lesson of the studies that have been conducted, namely that there is not a ‘safe’ messaging strategy that can always be relied upon to deliver the expected and intended effects.

---

25. 2021.

26. 2022b.

# The need to test messages

These findings suggest that practitioners should think hard about whether to conduct anti-corruption awareness raising campaigns, which may not offer value for money even if they do not backfire. If such campaigns are seen to be necessary – for example because under-acknowledging major social problems may give implicit license to a status quo that facilitates them – then it is essential to test their impact on a sample of the target audience *before* they are communicated to the wider public.

## Practitioners should think hard about whether to conduct anti-corruption awareness raising campaigns

Anti-corruption message testing should not rely exclusively on focus groups or interviews for two reasons. First, participants in such research may feel pressure to give researchers answers that participants think they want to hear, rather than report how they truly feel about a message. Second, anti-corruption messaging very likely impacts on attitudes subconsciously, and so individuals may not be fully aware of the true impact that a message is having.

Experimental techniques must therefore be used as they can provide a systematic estimate of the effect of exposure to messages. This strategy considerably reduces the risk of false reporting by individuals, as participants are never asked what a given message made them think or feel. This kind of pre-deployment testing should become a staple component of messaging development processes in any anti-corruption messaging campaign. Testing in this way while continuing to innovate how messages are framed and otherwise designed may also be the only way to find more effective messaging narratives that work as intended in the contexts they are deployed in. Further details on how such tests can be conducted can be found in Cheeseman and Peiffer.<sup>27</sup>

It is also important to note that even the most promising messaging campaign may be ill suited to change some behaviour. Citizens who come to be more critical about corruption because of an effective messaging strategy may be unable to do anything about it, for example, in highly authoritarian contexts where openly critical citizens are repressed. Moreover, in systems in which corruption is functional – i.e. it is

---

27. 2021; 2022.

necessary to get basic tasks done, such as overcoming economic, political and administrative blockages – refusing to engage in graft may be impractical for many ordinary people. This will be especially the case for those who have fewer resources and connections, and a large number of people who depend on them for food, education, and healthcare. Where both of these conditions hold, awareness raising efforts may not change behaviour even if it is successful in changing how people think about corruption.

*(For a more nuanced version of this argument that provides an example of how to test anti-corruption messaging and includes a longer discussion of core findings, see: [Cheeseman, N. and Peiffer, C. 2021. The curse of good intentions: Why anti-corruption messaging can encourage bribery.](#))*

# References

- Agerberg, M. 2021. [Messaging about corruption: The power of social norms](#). *Governance*, 35(3): 929-950.
- Byrne, S. and Hart, P. 2009. [The boomerang effect a synthesis of findings and a preliminary theoretical framework](#). (33)1: 3-37.
- Cheeseman, N., Peiffer, C. 2021. [The curse of good intentions: Why anticorruption messaging can encourage bribery](#), *American Political Science Review*. 116(3): 1081-1095.
- Cheeseman, N., Peiffer, C. 2022. [Why efforts to fight corruption can undermine the social contract: Lessons from a survey experiment in Nigeria](#). *Governance*. Early View.
- Cheeseman, N., Peiffer, C. 2022b. [Can messaging help us to fight SOC and corruption in Albania? SOC-ACE research paper, no. 2](#).
- Cheeseman, N., Peiffer, C. Forthcoming. [Anti-corruption awareness raising campaigns: Why do they fail, and how can “backfire” effects be avoided?](#) In Bratu, R. and Pozsgai-Alvarez, J. (ed)'s *Routledge Handbook of Anti-Corruption Research and Practice*. Routledge.
- Corbacho, A., Gingerich, D., Oliveros, V., Ruiz-Vega, M. 2016. [Corruption as a self-fulfilling prophecy: Evidence from a survey experiment in Costa Rica](#), *American Journal of Political Science*, 60(4):1077-92.
- Greszczuk, C. 2020. [Making messages work](#), The Health Foundation, 21 April 2020.
- Köbis, N., Troost, M., Brandt, C. and Soraperra, I. 2019. [Social norms of corruption in the field: Social nudges on posters can help to reduce bribery](#). *Behavioural Public Policy*. 6(4), 597-624: 1-28.
- Legros, S. and Cislighi, B. 2020. [Mapping the social-norms literature: An overview of reviews](#). *Perspectives on Psychological Science*. 15(1): 62-80.
- Paluck, E., Ball, L. 2010. [Social norms marketing aimed at gender-based violence: A literature review and critical assessment](#). *International Rescue Committee*.
- Park, H., Levine, T., Kingsley C., Orfgen, T., Foregger, S. 2007. [The effects of argument quality and involvement type on attitude formation and attitude change: A test of dual-process and social judgment predictions](#). *Human Communication Research*. 33(1): 81-102.
- Peiffer, C. 2017. [Getting the message: examining the intended – and unintended – impacts of corruption awareness-raising](#). *Developmental Leadership Program*.
- Peiffer, C. (2018). [Message Received? Experimental Findings on How Messages about Corruption Shape Perceptions](#). *British Journal of Political Science*, 50(3): 1207-1215.
- Peiffer, C., Walton, G. 2022. [Getting the \(right\) message across: how to encourage citizens to report corruption](#). *Development Policy Review*. 40(5): 1-23.
- Stead, M., Angus, K., Langley, T., Katikireddi, S.V., Hinds, K., Hilton, S., Lewis, S., Thomas, J.,



Campbell, M., Young, B., Bauld, L., 2019. [Mass media to communicate public health messages in six health topic areas: a systematic review and other reviews of the evidence](#). Public Health Research, PMID: 31046212.

# About the authors

## **Caryn Peiffer**

Caryn Peiffer is Senior Lecturer in International Public Policy and Governance at the School for Policy Studies at the University of Bristol. She has worked on a wide range of issues relating to anti-corruption, specialising in the impact of anti-corruption messaging, advising organisations such as Transparency International and OECD. Her studies have been conducted in a wide range of countries including Albania, Uganda, Indonesia, Botswana, Zambia, Papua New Guinea, Nigeria, and South Africa.

Twitter: @DrCarynPeiffer

## **Nic Cheeseman**

Nic Cheeseman is a Professor of Democracy at the University of Birmingham and was formerly the Director of the African Studies Centre at Oxford University. He mainly works on democracy, elections, and development and has conducted in-country research in a range of African countries including Ethiopia, Ghana, Kenya, Malawi, Nigeria, Uganda, Zambia, and Zimbabwe, but has also published on Latin America and post-communist Europe.

Twitter: @Fromagehomme

# Acknowledgements

Some of the research on which this analysis is based was funded by the Foreign, Commonwealth and Development Department (UK) through the SOAS-led Anti-Corruption Evidence (SOAS ACE) research consortium and the Serious Organised Crime and Anti-Corruption Evidence (SOC-ACE) research programme.

---

## Keywords

communication - anti-corruption strategy -  
social norms - citizen engagement

---

## How to cite

Peiffer, C.; Cheeseman, N.; (2023) Message misunderstood: Why raising awareness of corruption can backfire. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute (U4 Brief 2023:1)

### Publication

First published: 12 March 2023

### Disclaimer

All views in this text are the author(s)', and may differ from the U4 partner agencies' policies.

### Cover photo

Paul Keller (by) <https://www.flickr.com/photos/paulk/4608795889>



### Creative commons

This work is licenced under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence (CC BY-NC-ND 4.0)

---

## Partner agencies

German Corporation for International  
Cooperation – GIZ

German Federal Ministry for Economic  
Cooperation and Development – BMZ

Global Affairs Canada

Ministry for Foreign Affairs of Finland

Ministry of Foreign Affairs of Denmark /  
Danish International Development Assistance  
– Danida

Swedish International Development  
Cooperation Agency – Sida

Swiss Agency for Development and  
Cooperation – SDC

The Norwegian Agency for Development  
Cooperation – Norad

UK Aid – Foreign, Commonwealth &  
Development Office

Corruption erodes sustainable and inclusive development. It is both a political and technical challenge. The U4 Anti-Corruption Resource Centre (U4) works to understand and counter corruption worldwide.

U4 is part of the Chr. Michelsen Institute (CMI), an independent development research institute in Norway.

