



Working Paper 45

Strategic anti-corruption communications: Guidance for behaviour change interventions

Claudia Baez Camargo, Johanna Schönberg | June 2023

An INTEGRITAS Indonesia project publication

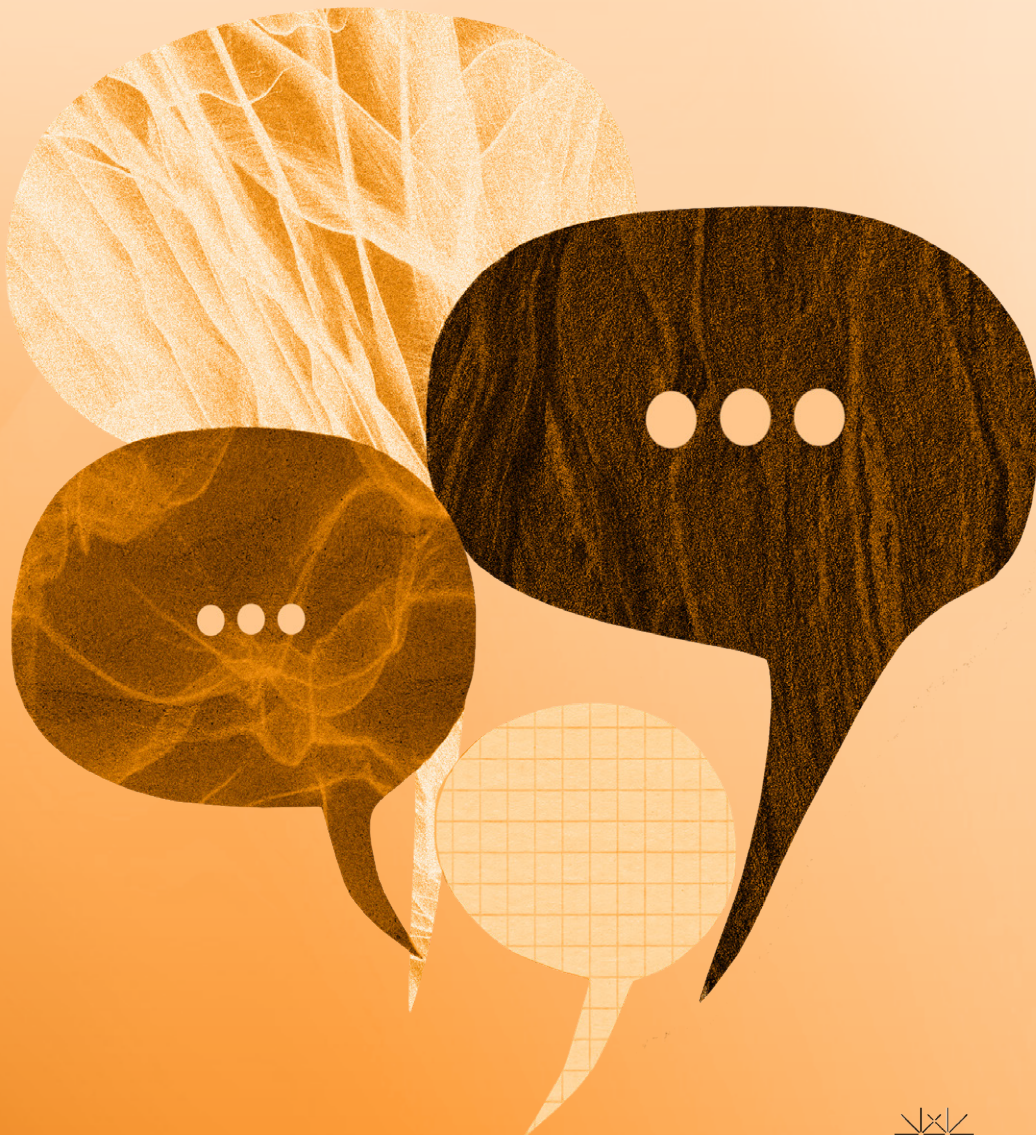


Table of contents

1 Introduction.....	9
2 An overview of the studies.....	10
2.1 Types of messaging	10
2.2 Negative or neutral intervention outcomes	11
2.3 Positive intervention outcomes	12
3 Findings.....	14
3.1 Emphasising that corruption is widespread typically backfires	14
3.2 Negatively framed messages are also mostly ineffective	15
3.3 More specific messages might be more effective	16
4 Anti-corruption messaging theory and concepts.....	18
4.1 Anti-corruption attitudes, behaviours and social norms	18
4.2 How can messages shape and change beliefs? Priming, learning and persuasion	20
4.3 Changing or updating beliefs: some insights from behavioural science	22
5 Theory of change.....	24
5.1 Developing a Theory of Change	24
5.2 Goals: What do we want to achieve with anti-corruption messaging?	25
5.3 Identify the specific problem and its context	26
5.4 Identify and formulate preconditions	27
5.5 Consider the baseline of the four behavioural categories	27
5.6 Consider the power of influence	28
5.7 Consider the messenger	28
6 Final recommendations	30
6.1 Invest in a strong monitoring and evaluation framework	30
6.2 Make the message relevant and actionable	31
6.3 Work with social norms	31
6.4 Develop a long-term strategic approach	32
7 Bibliography	33
8 Annex.....	35

About the authors



Claudia Baez Camargo
Head of Public Governance

Dr. Claudia Baez-Camargo is Head of Public Governance at the Basel Institute on Governance. In her role, she leads academic research and technical assistance programmes that seek to prevent and combat corruption in various countries around the globe. Claudia's work brings together academic research and technical assistance to promote context-sensitive anti-corruption approaches and address relevant corruption drivers. Among other areas, she has extensively researched how behavioural factors, such as those associated with social norms and mental models, can impact anti-corruption outcomes. She also works with a broad range of interested stakeholders on consultancy projects to develop context-sensitive strategies to prevent corruption in the public sector.

Claudia holds a PhD in Political Science from the University of Notre Dame, USA, and a graduate diploma in Economics from the University of Cambridge, England.



Johanna Schönberg
Public Governance Intern

Johanna Schönberg joined the Basel Institute in November 2022 as an intern with the Public Governance team. Prior to joining the Basel Institute, Johanna supported the British NGO Spotlight on Corruption and the Anti-Corruption and Integrity programme of the German development agency, GIZ. Johanna holds a master's degree in Corruption and Governance and a bachelor's degree in Philosophy, Politics and Economics from the University of Sussex, UK.

About this report

This publication is prepared as guidance for the USAID Indonesia Integrity Initiative (USAID INTEGRITAS), a corruption prevention programme in Indonesia implemented by a consortium consisting of Kemitraan (Partnership for Governance Reform), Indonesia Corruption Watch (ICW), Transparency International Indonesia (TII) and the Basel Institute on Governance. The initiative supports the country's efforts to prevent corruption by enhancing civic engagement and strengthening a culture of integrity and accountability.

The authors would like to thank Dr. Caryn Peiffer for comments provided to an early version of the manuscript.

This study is made possible by the support of the American people through the United States Agency for International Development (USAID). The contents are the sole responsibility of the Basel Institute on Governance and do not necessarily reflect the views of USAID or the United States Government.

Open-access licence and citation

The publication is part of the Basel Institute on Governance Working Paper Series, ISSN: . It is licensed under a Creative Commons Attribution NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

Suggested citation: Baez-Camargo, Claudia, and Johanna Schönberg. 2023. 'Strategic anti-corruption communications: a resource for practitioners.' Working Paper 45, Basel Institute on Governance. Available at: <https://baselgovernance.org/publications/wp-45>

Acronyms and abbreviations

BIT	Behavioural Insights Team
KPK	Corruption Eradication Commission (Indonesian Anti-Corruption Commission)
M&E	Monitoring and Evaluation
MAT	Medical Association of Tanzania
ToC	Theory of Change

Executive summary

This publication is intended to guide practitioners who are seeking to complement conventional anti-corruption measures by adopting a behavioural communications approach. It aims to connect a typology of anti-corruption messages with behavioural change theories, and discuss their impact. Subsequently, it suggests practical implications for designing anti-corruption communication as part of behaviour change interventions.

Seven key topically pertinent studies are reviewed to identify **typologies of messaging** and their **effectiveness**. The messages can be categorised as:

- Messages about the pervasiveness of corruption
- Messages about actions undertaken against corruption
- Messages about corruption as a moral issue
- Messages about the impact of corruption
- Messages about the illegality of corruption and sanctions
- Narrowly framed messages tailored to specific audiences

Overall, the review finds that:

- Emphasising that corruption is widespread typically backfires.
- Negatively framed messages are mostly ineffective.
- Messages with a narrowed focus and target group may be more effective.

Insights from behaviour change theories and research can provide further guidance:

- Social norms provide a helpful conceptual framework to tackle corruption as a collective action problem. Approaches that challenge social norms and reveal “true norms” have been found to be successful in fighting corruption.
- Messages can shape and change beliefs via priming (increasing awareness of and cognitive engagement with an issue), learning (new information leads to updating beliefs and perceptions), and persuasion (individuals are persuaded to think differently about the issue).
- Messages aimed at triggering learning and/or persuasion might be more effective if they are adequately contextualised, formulated and delivered.
- Certain messages fail to convince people to update their beliefs – even in the face of credible information that would appear to contradict held beliefs. Behavioural science suggests some approaches to overcome these challenges:

- People might be more open to updating their beliefs when their environment and/or welfare (external or internal) changes.
- Prompting people to seek additional information, rather than simply telling or informing them about relevant issues, may increase effectiveness.

Further, each anti-corruption communication intervention should be based on its own, tailored **Theory of Change (ToC)** to properly reflect the complexity of the targeted problem and the different activities and pathways that can ultimately lead to the desired change. The following stepping stones towards achieving such a ToC and suitable intervention entry points are detailed in this resource:

- Identify and formulate
 - specific goals;
 - the specific problem and its context;
 - the necessary preconditions which constitute the pathways for change.
- Consider
 - the baseline of prevailing descriptive norms, injunctive norms, personal attitudes and behaviours;
 - the power of influence (identify which outcomes are within the intervention's control);
 - the messenger (certain messengers hold more sway in reaching the target group than others).

Finally, the guidance recommends that practitioners:

- Invest in a strong **monitoring and evaluation framework**. This is crucial to measuring the achievement of outcomes effectively, and should be implemented continuously from the point of the intervention's development to its completion.
- Make the message **relevant and actionable**, by:
 - inviting relevant stakeholders to comment on the ToC logic and integrating their feedback into the ToC and intervention design;
 - observing real-world behaviour rather than merely surveying experiment results;
 - using messages about the concrete impact of corruption (particularly on the target's immediate environment) to stimulate the audience to take action.

- **Develop a long-term approach**, for instance by complementing messaging with:
 - changes in other factors of the choice architecture;
 - community meetings;
 - cultivating networks of champions;
 - engaging opinion leaders;
 - credible monitoring.

1 Introduction

Many anti-corruption initiatives contain some kind of messaging element, such as public education campaigns or awareness-raising activities. Substantive amounts of resources, financial and human, are invested every year to develop and deliver messages about the evils of corruption and the need to eradicate it. However, recent research has cast serious doubt on the effectiveness of messages about corruption and anti-corruption in achieving their intended results.

In fact, the evidence suggests that some of the most widely used formulations for anti-corruption messages either have no impact or, worse, backfire. There are promising results as well, but the framings of “successful” messages still need to be properly understood and probed for generalisability. It is therefore crucial to share the lessons learned from initiatives that have used and measured the impact of different types of anti-corruption messages and communicate them to practitioners. This will avoid wasting energy and resources on activities whose impact may be negligible or even negative. This urgency is reflected in recent publications (notably see [Peiffer and Cheeseman 2023](#)) that strive to synthesise the most important implications of this body of research.

This resource joins those efforts. It is intended to guide practitioners who might be seeking to complement, and even go beyond, conventional anti-corruption measures by adopting a behavioural communications approach. It seeks to connect a typology of messages to behavioural change theories, and discuss their impact. Subsequently, it suggests practical implications for designing anti-corruption communication as part of behaviour change interventions. This includes outlining how to develop a robust Theory of Change as a means to enhance the success of such efforts.

The guidance is based on a review of seven key topically pertinent studies that have been recently published. Anti-corruption practitioners who represent the target audience complemented the analysis with specific inputs. It must be noted, however, that the findings and recommendations presented are meant to provide practitioners with broad guidance, and not narrow prescriptions, on what needs to be kept in mind when developing strategic anti-corruption messages. It is not yet possible to make conclusive remarks or to generalise across contexts for several reasons.

First, many of the reviewed studies consisted of experiments on how people react to a survey which presents them with abstract ideas without much of the context that determines when and how corruption actually takes place in real life. Only two of the studies reviewed tested messages in an actual anti-bribery intervention in the field. People may respond differently when confronted with messages whose implications are more or less abstract (from abstract ideas about corruption to a corruption game) compared to when faced with instances of corruption in their real lives. Second, these experiments only tested changes after a one-off exposure to particular anti-corruption messages. More prolonged, regular and sustained exposure to messages may produce different outcomes.

2 An overview of the studies

The following studies were reviewed and will be referred to by their number and author name (e.g. Agerberg 1):

#	Study
1	Agerberg, Mattias (2021), Messaging about Corruption: the power of social norms, <i>Governance</i> , 35: 929-950.
2	Baez-Camargo, Claudia, Violette Gadenne, Dilhan Perera, Veronica Mkoji, Ruth Persian, Richard Sambaiga and Tobias Stark (2022). Using behavioural insights to reduce gift-giving in a Tanzanian public hospital: Findings from a mixed-methods evaluation. Basel Institute on Governance.
3	Cheeseman, Nic and Caryn Peiffer (2022), The Curse of Good Intentions: Why Anticorruption Messaging can Encourage Bribery, <i>American Political Science Review</i> , 116, 3: 1081-1095.
4	Köbis, Nils C., Marleen Troost, Cyril O. Brandt and Ivan Soraperra (2019) Social Norms of Corruption In The Field: Social Nudges On Posters Can Help To Reduce Bribery, <i>Behavioural Public Policy</i> : 1-28
5	Peiffer, Caryn (2018), Message Received? Experimental Findings on How Messages about Corruption Shape Perceptions, <i>British Journal of Political Science</i> , 50: 1207-1215.
6	Peiffer, Caryn and Grant W. Walton (2022) Getting the (Right) Message Across: How to Encourage Citizens to Report Corruption, <i>Development Policy Review</i> : 1-23.
7	Sharot, Tali, Max Rollwage, Cass R. Sunstein and Stephen M. Fleming (2022), Why and When Beliefs Change, <i>Perspectives on Psychological Science</i> : 1-10

2.1 Types of messaging

The seven studies present various types of anti-corruption messages, differing in content, framing and target. They can be outlined as follows:

Type of message	Summary of content
Messages about the pervasiveness of corruption	<ul style="list-style-type: none"> – Prevalence of grand corruption (Peiffer 5) – High levels of petty corruption (Peiffer 5) – Corruption is widespread in society (Peiffer and Walton 6, Cheeseman and Peiffer 3) – Contradicting injunctive social norms about the acceptability of corruption by others (Agerberg 1) – Highlighting a decreased prevalence of petty corruption (challenging perceptions on descriptive social norms) (Köbis et al. 4)

Type of message	Summary of content
Messages about actions undertaken against corruption	<ul style="list-style-type: none"> – Government success (Peiffer 5, Cheeseman and Peiffer 3) – Citizen engagement and action (Peiffer 5)
Messages about corruption as a moral issue	<ul style="list-style-type: none"> – Role of religious leaders (condemning corruption) (Peiffer and Walton 6, Cheeseman and Peiffer 3) – Reminder of professional ethics, appealing to professional identity (Baez-Camargo et al. 2)
Messages about the impact of corruption	<ul style="list-style-type: none"> – On local communities (Peiffer and Walton 6, Cheeseman and Peiffer 3)
Messages about the illegality of corruption and sanctions	<ul style="list-style-type: none"> – Corruption is illegal (Peiffer and Walton 6) – Bribery is punishable (Baez-Camargo et al. 2) – Alluding to taxes (Cheeseman and Peiffer 3)
Narrowly framed messages tailored to specific audiences	<ul style="list-style-type: none"> – Health workers emphasising professional ethics (Baez-Camargo et al. 2) – Users of health facilities emphasising that gift-giving (euphemism for bribery) is corruption and is forbidden in the health facility (Baez-Camargo et al. 2)

2.2 Negative or neutral intervention outcomes

Peiffer's (5) study in Jakarta tested four messages regarding either pervasiveness of corruption or actions undertaken against corruption (prevalence of grand corruption, prevalence of petty corruption, government success and citizen engagement), as outlined in the table above. The study found that all four messages:

- Increased the degree of worry about the impact of corruption on development.
- Decreased pride in the government's anti-corruption efforts.
- Decreased confidence in that it is easy to report and get involved in the fight against corruption.
- Did not alter perceptions about levels of corruption.
- The petty corruption message caused increased scepticism about the government's efforts to fight corruption.
- Overall, all messages seemed to contribute to corruption fatigue.

Cheeseman and Peiffer (3), using a bribery game, looked at the effect of five different anti-corruption messages (on corruption pervasiveness, anti-corruption action, corruption as a moral issue, impact of corruption and illegality of corruption) on behaviours in Lagos, Nigeria, and found that:

- The messages failed to discourage corrupt behaviours.
- The messages about widespread corruption, religious leaders and government success even made individuals more willing to bribe.
- However, pre-existing perceptions matter: messages backfired most strongly with those who were already pessimistic about the prevalence of corruption. For non-pessimists, four messages did not trigger a significant effect; however, the message on the illegality of corruption, alluding to taxes, had a positive effect (see section 2.3).

Additional evidence from other studies points to adverse or neutral outcomes of anti-corruption messaging:

- Chong et al. (2015) found that exposure to information about a mayor's corrupt deals did not change perceptions of the level of corruption.
- Corbacho et al. (2016) conducted a survey experiment in Costa Rica, finding that a message about an increasing number of people witnessing corruption made respondents more willing to pay a bribe.
- Falisse and Leszczynska (2021) found that an anti-corruption sensitisation message on professional values of integrity in public service delivery did not change bribe-taking, nor did a general message on the importance of good governance.

2.3 Positive intervention outcomes

Peiffer and Walton's (6) study in Papua New Guinea also tested four messages and found that:

- The message highlighting the impact of corruption on local communities improved the willingness to denounce corruption because it is the morally right thing to do - even if it involves spending a day in court.
- The other messages (referring to anti-corruption as a moral duty, corruption's illegality and pervasiveness) did not change the willingness to report.

Cheeseman and Peiffer (3) and found that:

- Non-pessimists' exposure to the message alluding to taxes – which implied that “corruption represents the theft of taxes and fees paid by ordinary citizens” and thus impacts the public services (ibid., p.1806) – was found to significantly decrease their probability of bribing.

Agerberg (1) found that:

- Those exposed to an injunctive norms “corrective” message (e.g. revealing that most people in their context actually do not condone corruption contrary to conventional wisdom):
 - Had more trust in others.

- Were less inclined to respond that corruption is part of their country's culture.
- Showed a lower willingness to pay a bribe.

Köbis et al. (4) found that:

- Exposure to posters indicating decreased bribery in a community:
 - Reduced the perceived prevalence of bribery.
 - Reduced the offering and accepting of bribes in a bribery game.

Baez Camargo et al. (2) found that:

- Posters and desk signs conveying anti-bribery messages in a public health facility significantly decreased the intention to offer a bribe among users.
- Anti-bribery messages aimed at health workers, appealing to their professional ethics, positively changed their attitudes towards soliciting and receiving bribes.

3 Findings

3.1 Emphasising that corruption is widespread typically backfires

Messages that inform that corruption is widespread appear to contribute to corruption fatigue. This finding is corroborated across contexts. One study suggests that “perception of widespread corruption is associated with an unwillingness to protest corruption, report corruption to the authorities, and pay more for a product produced by a company that has not engaged in corruption” (Peiffer 5: 1208). Part of the reason why messages about the pervasiveness of corruption backfire is that they do not provide any new useful information. When people are informed that corruption is widespread, this may simply validate beliefs about the inevitability of corruption. In other words, such messages confirm the expectation that most other people are partaking in acts of corruption, as appears to be the case in Indonesia (see Figure 1).¹

Messages emphasising that corruption is widespread give information about what people believe most other people are doing (the descriptive social norm). We know that people can often be prone to following descriptive social norms (Bicchieri, 2017). This might explain why, in some studies, such messaging is associated with an increase in the propensity to bribe and, hence, why information campaigns highlighting the corrupt deals of political figures do not seem to have a positive effect and can often backfire.

Perceptions of corruption in Indonesia

In Indonesia, a majority of people believe corruption to be widespread according to data from the World Values Survey:

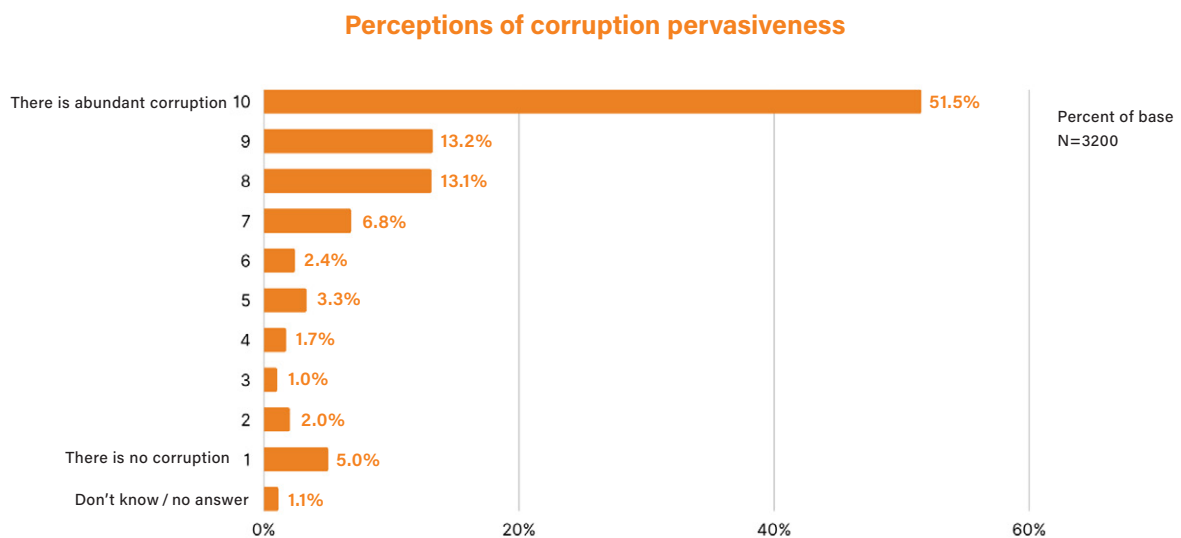


Figure 1: Perceptions of corruption in Indonesia. Source: *World Values Survey (Dataset 2018, Indonesia)*

¹ This is furthermore consistent with the findings from a survey conducted by the Basel Institute's Green Corruption team in Indonesia in 2021, which found that the share of the public that says corruption has increased over the last two years has been rising (Grossman et al 2021).

3.2 Negatively framed messages are also mostly ineffective

Strictly normative framings – highlighting that corruption is wrong, illegal, immoral or unethical – also do not appear to work. This is probably linked to the fact that, as numerous empirical observations suggest, people tend to strongly condemn corrupt practices, even in contexts where corruption levels are high. In Indonesia, a majority of people strongly condemn corrupt practices (bribery and claiming government benefits one is not entitled to) as suggested by data from the World Values Survey (see Figures 2 and 3).

Justifiable: Someone accepting a bribe in the course of their duties

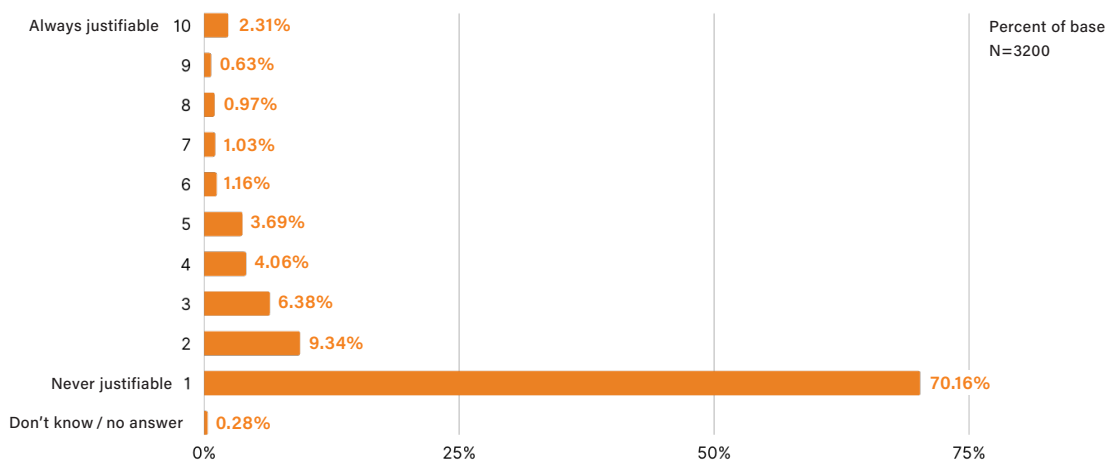


Figure 2: Justifiability of someone accepting a bribe in the course of their duties. Source: World Values Survey (Dataset 2018, Indonesia)

Justifiable: Claiming government benefits to which you are not entitled

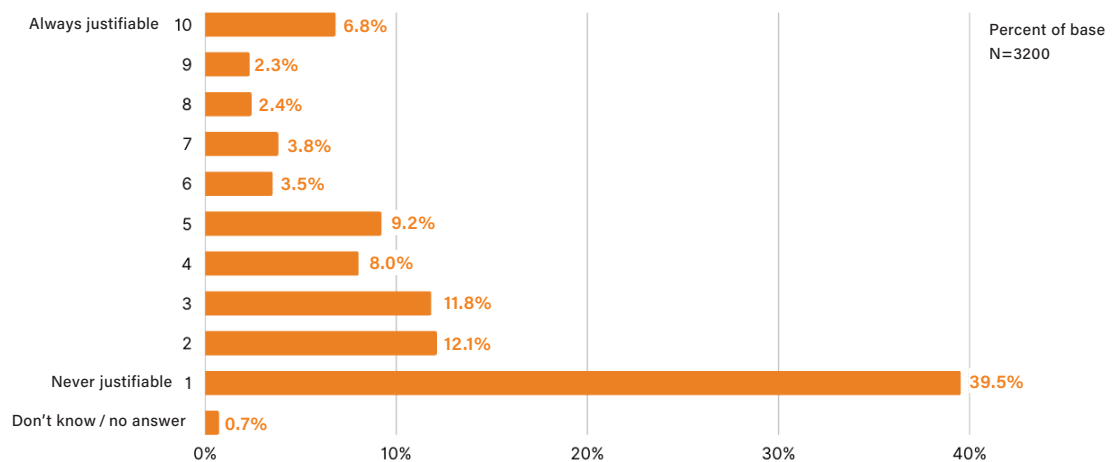


Figure 3: Justifiability of claiming government benefits to which you are not entitled. Source: World Values Survey (Dataset 2018, Indonesia)

Therefore, messages that aim to raise awareness about the evils of corruption are not providing new or valuable information.

Thus, individuals might feel that, although perhaps personally repulsed by it, corruption is inevitable and inescapable. They might also believe that people around them view corruption as an acceptable behaviour and engage in it regularly. Such a situation can be, understandably, linked to corruption fatigue, whereby people might become cynical or apathetic, or even give in to partaking in acts of corruption.

3.3 More specific messages might be more effective

As indicated above, generic messages about corruption being widespread and raising awareness about the evils of corruption are plausibly linked with corruption fatigue. That notwithstanding, the reviewed studies also found positive results which suggest that giving messages a narrower focus might be more effective.

One aspect of narrowing down involves targeting particular audiences and tailoring messages accordingly. An important distinction uncovered by several of the studies reviewed is that between people who already hold pessimistic views about corruption in their contexts and others who have more positive outlooks vis-à-vis the inevitability of corruption to begin with:

- The study by Cheeseman and Peiffer (3) suggests that people who already hold negative views, especially those who have personally encountered corruption, tend to react more negatively to anti-corruption messages, especially to those that underscore the prevalence of corruption. In this study, this group became more willing to engage in bribery.
- The same study also found that those with a more positive outlook became less willing to engage in bribery when exposed to a message about taxation.
- In the study conducted in Mexico (Agerberg 1), those with the most negative prior perceptions reacted more strongly (positively) to the message.
- The implications stemming from these three examples are, evidently, inconclusive. Very preliminarily, they would seem to suggest that:
 - People with negative views about the prevalence and perspectives of dealing with corruption should not be targeted with messages that reinforce that narrative.
 - This audience could instead be targeted with messages conveying credible information that challenges or contravenes conventional wisdom, as suggested by the Mexico study (Agerberg 1). However, Cheeseman and Peiffer's study (3) invites caution. In the Nigerian case, even positive messages (for example, about leaders' political and religious commitment to fighting corruption) triggered an increased willingness to bribe.

Another way to narrow the messages' focus is to target concrete types or patterns of corruption. This was the approach Baez Camargo et al. (2) took in their intervention, which addressed bribery in a Tanzanian hospital. By narrowly specifying the intervention target, the main target groups (health workers and users of public health services) can easily be identified and messages can be tailored for each group. The messages were thoroughly localised, for example, by ensuring the language used to refer to the targeted behaviour (exchange of gifts with the intention to win special favours) matched that used by those involved in it. The study's results indicated that this method was effective.

4 Anti-corruption messaging theory and concepts

This section extracts theoretical insights from the studies reviewed and connects them with the messaging typologies presented above.

4.1 Anti-corruption attitudes, behaviours and social norms

Social norms have recently become a topic of interest for anti-corruption practitioners. The resilience of specific patterns of corruption despite legal and institutional reforms has shifted the attention to problems of collective action, and social norms provide a conceptual framework that is increasingly supported by empirical evidence from various countries.²

Social norms are loosely defined as unwritten rules and beliefs about what is considered acceptable in a particular social group. In academic literature, the distinction is made between descriptive and injunctive social norms:

- **Descriptive social norms** are norms based on the perceived dominant patterns of behaviour within a given group (what one believes most people are doing) in response to a particular situation.
- **Injunctive social norms** are norms about the perceived acceptability of a particular behaviour (what one believes most people think is considered the right course of action in response to a particular situation).

Social norms might come into play when people perceive:

- That "everybody is engaging in corruption" (descriptive social norm) – *"If everybody bribes and I am the only one who doesn't, then I lose out. Therefore, I bribe even if it goes against my personal beliefs."*
- That "corruption is the accepted and expected behaviour and not partaking will be criticised or otherwise socially punished" (injunctive social norm) – *"I am a public official, and my family and friends expect me to use my position to gather resources and benefits for them. If I don't do so, they will criticise me and maybe even ostracise me."*

Corruption and social norms can lead to collective action dilemmas when people find themselves socially pressured to engage in corruption despite their personal preferences, or when perceptions about social norms are influenced by stereotypes or biased information and thus not aligned with true collective preferences.

2 Collective action dilemmas are a category of social problems identified in political philosophy and economics (very prominently in game theory) to refer to situations where individuals would be better off if they worked together to achieve common goals in line with their individual interests but nonetheless fail to do so (see the classical work by Mancur Olson, *The Logic of Collective Action*, 1965). Corruption can sometimes be understood as a collective action problem, as famously laid out by Persson, Rothstein and Teorell (2013). This entails that while actors might comprehend that eliminating corruption would be advantageous to them, "they cannot trust that most other actors will refrain from corrupt practices", and, as such, "they have no reason to refrain from paying or demanding bribes" (ibid., p.457). Therefore, the "shared expectations about other actors' behavior" (ibid.) typically undermine conventional anti-corruption incentives. These "shared expectations" fall under the category of 'descriptive social norms', as will be explained in this section.

Based on the above, research teams have developed and implemented approaches to reveal “true norms” to fight corruption. There are three main versions of these approaches:

Version 1: Revealing the true injunctive norm

This is what Agerberg (1) tried to address in his study. The goal was to correct perceptions about the injunctive social norm.

In Mexico, people share the conventional wisdom that corruption is socially accepted and, even more, validated as a legitimate (often the only) route to ascend the socioeconomic ladder or make a decent living wage. However, survey results show that, in reality, most Mexicans strongly oppose and morally condemn corruption. Agerberg's study involved revealing the true injunctive norm that, in fact, most Mexicans reject and condemn corruption. This yielded positive results.

Version 2: Revealing the true descriptive norm

There is evidence to suggest that people consistently overestimate the prevalence of corruption: measurements about perceptions of the prevalence of corruption almost invariably yield much higher values than measurements of actual experiences with corruption. Challenging this narrative might be helpful, as the study by Köbis et al. (4) demonstrated. This intervention elicited positive results by placing posters in a South African city indicating that increasingly more people rejected bribery.

These approaches are both promising but, of course, need to be based on true facts. That is, the credibility of the information used appears to be of utmost importance.

Version 3: Challenging both accepted descriptive and injunctive norms

Baez Camargo (2) and team devised an intervention to tackle both descriptive and injunctive norms of bribery in Tanzanian health facilities. The approach was two-pronged and consisted of the following components:

1. Devising a message appealing to the professional ethics of health workers. The message was introduced to all staff in the treatment departments in a hospital via a letter coming from the Medical Association of Tanzania (MAT) and the hospital management. In addition, a group of hospital staff were recruited as “champions” to disseminate the anti-corruption message through their social networks amongst their peers. The idea was that if the message came from both authoritative figures and trusted individuals in the hospital staff's social networks, the message would indicate a shift in the injunctive social norm (i.e. accepting “gifts” from users is no longer socially acceptable).
2. Placing posters and desk signs conveying the anti-bribery message around the treatment hospital units targeting primarily the facility users but also altering the hospital environment or the so-called “choice architecture.” This part of the intervention was meant to signal a change in the descriptive norm (i.e. most people are no longer offering “gifts” to health workers).

This intervention achieved an 18% reduction of 18 percentage points (from 43 percent to 25 percent) in the future intention to offer a bribe on the part of the hospital users, and it was also commended by the hospital staff.

4.2 How can messages shape and change beliefs? Priming, learning and persuasion

According to Peiffer (5) we can identify three different mechanisms by means of which anti-corruption messages can help shape and change perceptions:

1. **Priming:** A message makes the individuals become more aware and think more about the issue that would have otherwise been the case.
2. **Learning:** New information leads to updating of perceptions and beliefs.
3. **Persuasion:** Individuals are persuaded to think differently about the issue (e.g. by appealing to emotions, or because the messenger is influential).

Several of the studies reviewed suggest that priming may lead to or reinforce corruption fatigue and, therefore, be counterproductive when it underscores information about the prevalence and the evils of corruption that is already known to people and confirms their pre-existing beliefs. This approach seems to backfire, particularly when people are already pessimistic about corruption.

Messages aimed at triggering learning and/or persuasion might be more effective if they are adequately:

1. **Contextualised:**
 - Tailored to the particular way in which the targeted practice of corruption happens in its context.
 - Aligned with reality and supported by credible evidence.
2. **Formulated:**
 - Using the right vocabulary that will be familiar and clearly linked to the targeted behaviour by the intended intervention target groups.
 - Striking the right tone (positive, empowering, threatening) that will attract attention and elicit a reflection conducive to updating one's beliefs.
 - Addressing the causes of corruption fatigue, for example by challenging conventional wisdoms.
3. **Delivered:**
 - The messenger matters – the source of the information should be regarded as trustworthy by the members of the target group.
 - Different dissemination channels might be needed to reach different audiences.

However, the specific pathways to change following these dimensions remain unclear. The findings from the studies reviewed are ambivalent and should still be considered with caution.

Agerberg (1) elicited positive results by informing people that the situation is not as bad as they thought in terms of the attitudes towards corruption amongst fellow citizens in Mexico. This “positive information shock” provided new information that challenged conventional wisdom using a positive tone and, in doing so, succeeded in changing beliefs about injunctive social norms concerning bribery in Mexico. A similar finding was observed in the study by Köbis et al. (4), where posters communicated novel information contravening stereotypes in a positive tone regarding reduced levels of bribery. These two studies would appear to show that new, positive information triggers positive results.

However, other studies that used a positive tone and provided new information did not achieve a similar positive outcome. Peiffer’s (5) study in Indonesia tested two positively framed messages about the success of the Corruption Eradication Commission (KPK), Indonesia’s anti-corruption agency, in fighting corruption and the ease of civic engagement in the fight against corruption. However, they failed to elicit a positive outcome and rather decreased pride in the government’s efforts and reduced the extent to which people thought it easy to engage in supporting anti-corruption goals.

A similar ambivalent situation can be found where messages appeal to persuasion, for example by appealing to the role of leaders and authoritative figures as opinion makers. Baez Camargo et al (2) framed their intervention with the premise that the messenger matters and had the anti-bribery message delivered to health workers by their professional association and hospital manager as well as their own peers with positive results. However, in the case of Cheeseman and Peiffer (3) their message emphasising the role and commitment of political and religious leaders in the fight against corruption backfired. Similarly, in Peiffer and Walton’s study in Papua New Guinea (6), the message underscoring the role of religious leaders in fighting corruption failed to change respondents’ attitudes towards reporting corruption.

The difference might be that in the study by Baez Camargo et al. (2) the message was actually delivered by the actors who were deemed to be in a position to play the role of trendsetters, whereas in the other studies, the role of the trendsetters was merely mentioned in a survey. More research is needed in order to arrive at more robust conclusions in this regard.

Persuasion can also be triggered by appealing to affective elements, such as the impact of corruption on family and the own community. Peiffer and Walton’s (6) findings in Papua New Guinea provide positive evidence in this respect although the relevance of this framing was not replicated in Cheeseman and Peiffer’s (3) study in Nigeria.

4.3 Changing or updating beliefs: some insights from behavioural science

A number of the studies reviewed suggest that some messages fail to convince people, especially corruption pessimists, to update their beliefs. They fail even in the face of factually based, credible information about positive trends or other events that would appear to contradict beliefs already held.

Behavioural science might offer some insights as to why this might be the case and suggests some approaches to overcome these challenges. The following is based on Sharot et al. (7).

These authors have developed a framework to understand when and why people change their beliefs. The framework distinguishes four dimensions:

- **External vs. internal outcomes:** These refer to the benefits (or utility) people obtain from holding certain beliefs. These might be external, when a concrete benefit is accrued from holding one particular belief (e.g. people holding certain political views might be favoured for obtaining public employment). They might also be internal (e.g. people enjoying a sense of peace or happiness for holding certain religious beliefs).
- **Accuracy dependent vs. independent:** Holding a belief might be contingent on whether it is empirically verifiable or not. For example, some vaccine sceptics may change their beliefs when confronted with evidence about vaccine safety and effectiveness, as the updated belief has higher utility to them. But others will not, as the cost of altering their beliefs may be felt to be too high.

The authors suggest that beliefs are contingent on what combination of the above provides each individual the greatest utility (or value).

This framework accounts for why there are some cases in which people will not change their beliefs even when new, credible evidence to the contrary is presented to them. For example, individuals fail to update their beliefs in the face of information pointing towards unpleasant conclusions.

An interesting insight presented by these authors is that people might be more open to updating their beliefs when their environment changes. This could suggest that the timing of anti-corruption messaging might be as important as the formulation and content of the messages themselves. Windows of opportunity might open, for example, during periods immediately after an election or when a significant event has changed perceptions about the level of threat or emergency (such as the uncovering of a massive corruption scandal or even a natural disaster).

Furthermore, it seems important to prompt people to seek additional information rather than simply telling or informing them about relevant issues (raising awareness). According to Sharot et al. (7), the research indicates that when people are less confident about their beliefs, they will be more likely to seek new information. In this sense, challenging conventional wisdoms might be an important approach, but is only the start to trigger a desired response.

The study by Sharot et al. (7) does suggest that even though individuals might hold beliefs that are not aligned with empirical facts, they will revise and update their beliefs when it affects their welfare (either external or internal) in a noticeable manner. Given that we know the hugely detrimental effects that corruption has on development generally and on citizens' welfare in particular, this insight provides hope that it should be possible to identify entry points for designing interventions that more compellingly incentivise people to act in support of the fight against corruption.

5 Theory of change

5.1 Developing a Theory of Change

Having a Theory of Change (ToC) is crucial for any intervention. Nonetheless, this is only sometimes contemplated or explicitly formulated.

A commonly held belief (which in itself could be understood as comprising a ToC of sorts) is that by raising awareness about corruption (prevalence, consequences, illegality thereof), citizens will be inspired, outraged or in some way triggered into springing into action. However, according to the reviewed literature, such an awareness-raising approach is not enough to incite citizens to act against corruption and certainly not immediately. Therefore, each anti-corruption communication intervention should be based on its own, tailored ToC to properly reflect the complexity of the targeted problem and the different activities and pathways that can ultimately lead to the desired change.

Difference between results chain and ToC

- Results chain: shows a causal logic, that an input (e.g. an activity) leads to an output. If X occurs, then Y follows.
- ToC: builds on the results chain, and tests its causal logic by making explicit assumptions about how the input leads to the output, i.e. “about how change happens”; if X occurs, then Y follows, because necessary preconditions have been fulfilled (Johnson, 2012, p.9).

The ToC needs to reflect change along a continuum – where encouraging changes in perception is a primary step towards evoking behavioural outcomes. The following ideas might help stimulate the development of Theories of Change to inform strategic anti-corruption communication efforts. These are meant to inspire but in no way substitute for an adequate understanding of the characteristics of each context and target audience in line with the goals of each project. We strongly recommend practitioners refer to USAID’s Theory of Change guidance (Salib, 2022), which provides an overview of important elements and questions to consider while developing a ToC.³

³ See Annex for an example of a detailed ToC for a hospital intervention (Baez-Camargo et al., 2022).

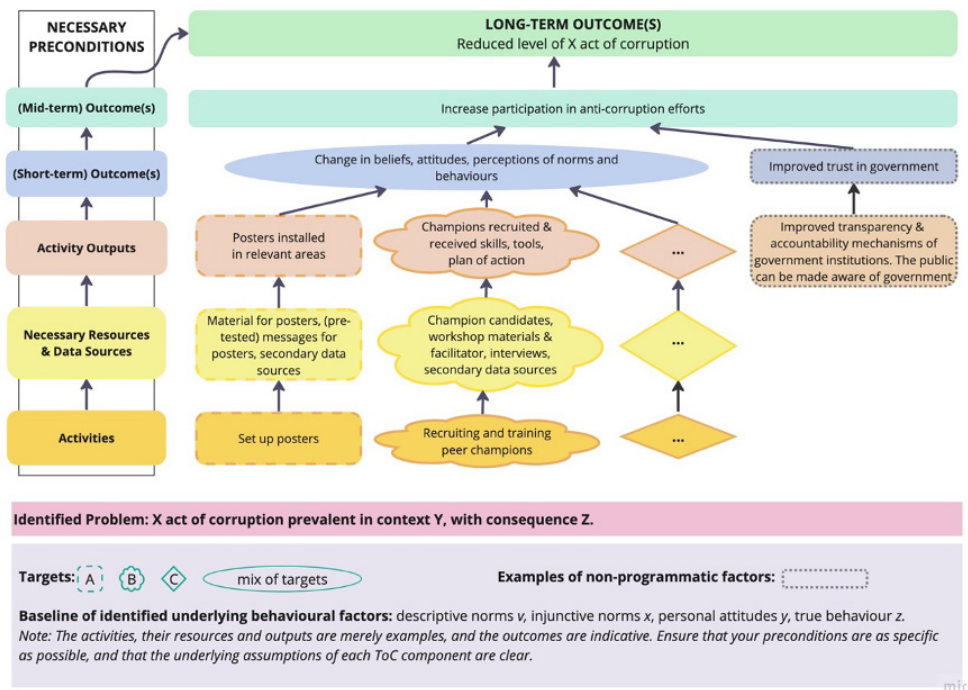


Figure 4: Theory of Change template

5.2 Goals: What do we want to achieve with anti-corruption messaging?

A point of departure for developing any anti-corruption intervention is clarity about the concrete results the strategic anti-corruption communication campaign is trying to achieve. Although the ultimate goal of many interventions is to change behaviours associated with negative development outcomes, changes in attitudes and beliefs are often a prerequisite.

The following are examples of indicative results that might be sought through different types of anti-corruption messaging:

Goal: Changing perceptions and/or beliefs about:

- Level of corruption:
 - Prevalence: among public officials, in sectors and generally
 - Changes through time: increase, decrease
- Impact of corruption:
 - Grand corruption
 - Petty corruption
- Government’s actions to fight corruption:
 - Credibility of government’s anti-corruption actions
 - Pride in government’s actions

- Ability of citizens to act against corruption
- Other people's beliefs and behaviours (revealing true preferences, i.e. when perceptions about the descriptive norm are not aligned with people's preferences)

Goal: Encourage citizens to act against corruption (behaviour change) through:

- Monitoring corruption (as tested by Peiffer and Walton (6))
- Filing reports of corruption instances
- Advocacy
- Voting

Goal: Compel people to refrain from participating in acts of corruption (behaviour change) with the effect of:

- Reducing bribery
- Reducing favouritism
- Reducing vote-buying attempts
- Increasing law abidance

5.3 Identify the specific problem and its context

To increase the rate of an intervention's success and ensure efficient resource allocation, the specific problem that is being targeted and its context must be identified and operationalised. Only on this basis can the ToC be sufficiently specific so as to find appropriate entry points.

For instance, *"corruption is rampant in healthcare"* – is a general problem, a description of a general situation. A more specific problem articulation would be: *"Bribes are frequently exchanged between hospital staff and users in hospital(s) xyz, impeding fair access to healthcare and endangering patients' health."*

Context matters – behaviour can only be changed where the wider framework allows it. For instance, Cheeseman and Peiffer (3) point out that behaviour cannot be changed where corruption is *"functional"*.⁴ In such situations, *"refusing to engage in graft may be impractical for many ordinary people"* (Peiffer and Cheeseman, 2023, p.15). Thus, the choice architecture (including laws, regulations and informal incentives) as well as structural issues (low wages, scarcity of medicines or medical supplies) must also be addressed to deal with the functional role of corruption.

Furthermore, in some cases, refusing to give in to extortive acts of corruption can have deleterious consequences. For instance, when a pregnant patient is

⁴ 'Functional' corruption is where corruption is "necessary to get basic tasks done, such as overcoming economic, political and administrative blockages" (Peiffer and Cheeseman, 2023, p.14).

demanded to pay a bribe in exchange for the service of delivering her child (albeit being covered by health insurance or having paid fees), refusing to pay the bribe can lead to severe complications. Therefore, not only the context of the bribe-demanding party must be changed, but a proper risk assessment of the possible unintended consequences of fighting corruption through a “*do no harm*” lens must be carefully considered.

5.4 Identify and formulate preconditions

Once the goal has been clearly set, and the specific problem has been formulated, the pathways for change must be identified. These consist of specific necessary preconditions and initial outcomes that need to be achieved in order to reach the desired ultimate outcome.

Questions to be considered when identifying both the ultimate outcome and preconditions include the following (USAID; Johnson, 2012):

- Which outcomes (ultimate outcome and preconditions) are feasible considering the local context and possible entry points?
- Who is the target group?
- Are we formulating concrete changes, not just project deliverables?
- On what level (i.e., system level, local level, individual level) are we aiming for changes?
- What type of behavioural change must take place to achieve the ultimate outcome?
- Are the ultimate outcome and the preconditions SMART (specific, measurable, achievable, realistic, and time-bound)?
- Can the outcomes be solely achieved by the programme? What is in the programme’s power of influence?
- What resources are needed as preconditions?

Once the preconditions have been formulated, as many obstacles as possible must be removed from the pathways of change; interventions must be suggested, tackling each obstacle and programmatic precondition. Thus, the ToC is to be developed in a manner that makes the expected causal chain clear, explaining how change is expected to come about in incremental steps.

5.5 Consider the baseline of the four behavioural categories

There are four categories of underlying assumptions and expectations (“behavioural categories”) that need to be considered: prevailing descriptive norms, injunctive norms, personal attitudes and behaviours.

We do not know a priori how changes in the perceived norms and personal beliefs may interact in order to yield a change in behaviour (which is the ultimate goal).

For instance, a study in Rwanda analysed attitudes towards ethnicity (after the genocide), and how these could be changed using an “edutainment” approach via soap operas (Paluck, 2009). The study observed positive behaviour change in the form of more inter-ethnic marriages and less hostility in the social context. In comparing the baseline and endline data of the behavioural categories, the researchers found that, while the edutainment campaign had worked to change perceptions about the social norms (e.g. interethnic violence and xenophobia are not acceptable), the target group’s personal beliefs had not changed. Other studies, however, propose that personal beliefs must change first, to then elicit changes in social norms and behaviours (see for example Bicchieri, 2017).

Hence, quantitative baseline and endline data on the four behavioural categories should be collected and analysed to track how changes are being elicited by the intervention. Depending on the intervention design, it may also be necessary to conduct qualitative interviews to better trace causality. Only in this way can practitioners and researchers systematically contribute to the development of evidence-based ToCs.

5.6 Consider the power of influence

As mentioned above, we need to consider what our power of influence is when developing a ToC; what is and is not in the intervention’s control. This includes distinguishing between programmatic and non-programmatic preconditions and assessing the choice of the messenger.

Programmatic preconditions are addressed by the programme itself or other actors. For instance, a programmatic precondition would be to increase awareness of corruption implications and thus increase public participation in anti-corruption efforts. Meanwhile, non-programmatic preconditions cannot be addressed by the intervention. Still, they should be monitored so that when the context changes, the programme can be adapted to exploit new windows of opportunity. For example, it is not in our power of influence to improve the transparency and accountability of government institutions. However, it is a precondition to improve trust and confidence in the government, which in turn would be a precondition to increase public participation in anti-corruption efforts.

5.7 Consider the messenger

The power of influence also lies in the choice of messenger: certain messengers hold more sway in reaching the target group than others. This impact of the choice of messenger is called the “messenger effect” (Maclean, Buckell, and Marti, 2019); the messenger choice impacts the “level of influence” of the message being delivered.

Engagement is higher when the messenger is trusted, i.e., a person regarded as a “credible source” by the target audience. Such people can be peers or persons of authority, such as agents with certain expertise (either via personal experience or scientific knowledge), “community influencers” or celebrities.

A helpful resource on selecting “good” messengers is the Behavioural Insights Team (BIT) and TRAFFIC’s joint guideline, which goes into further detail regarding the effectiveness of choosing varying persons of authority and peers (TRAFFIC and BIT, 2019).

Credibility matters for good messenger selection (BIT & TRAFFIC, 2019)

- Perceived credibility depends on the audience; different audiences may have a different perceived credibility of the messenger in question.
- 3 characteristics of credibility:
 - Inclination towards truth
 - Presentation of truth
 - Potential of truth
- Further relevant questions: Will the audience trust the motive? Is the messenger liked by the audience?

The study in Tanzania (Baez-Camargo et al. 2) chose staff champions within peer networks as the messengers, along with the Medical Association of Tanzania (MAT) and hospital management to formally support the programme. The study's results confirmed the overall positive impact of using a peer-led approach. Furthermore, relationship closeness of colleagues was found to positively impact the propensity to talk about the intervention and gift giving (ibid.).

Overall, health workers reported that they felt the dual approach of utilising peer networks and the endorsement of the MAT board was "effective" and "the support of the message by formal authority figures was considered to add credibility to the intervention" (Claudia Baez-Camargo, 2022, p.21).

6 Final recommendations

6.1 Invest in a strong monitoring and evaluation framework

A well-thought-out and implemented monitoring and evaluation (M&E) framework is crucial to measure the achievement of outcomes effectively. An M&E framework should target all the components of the developed ToC. This is needed to be able to compare the effects of the intervention with its objectives and is helpful in identifying factors that contribute to or constrain outcome achievement.

M&E is a continuous process that should accompany the intervention from the point of development to its completion. As such, results from M&E can contribute to developing and understanding the ToC and can be linked with testing messages before rolling out the communication campaign. Indeed, M&E must be implemented continuously to ensure that:

- the intervention's incremental adaptability to changes (e.g., due to change in context, new insights, in response to any issues arising during the intervention activities);
- the intervention has the most significant possible impact;
- sufficient evidence is produced on the intervention's attributable results, strengths and weaknesses.

To adequately track change, ToC concepts must be operationalised into measurable indicators (key performance indicators). Indicators should answer the question: *"How will we know that outcome x has been achieved?"* The indicators should provide the value used to measure the implementation and efficacy of the communication intervention.

Useful resources to consult in the process of developing an M&E framework, and identifying key indicators, are:

- Wathne, Cecile (2022) Effectively evaluating anti-corruption interventions. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute (U4 Issue 2022:6)
- Johnsen, Jesper (2012) Theories of change in anti-corruption work: A tool for programme design and evaluation. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute (U4 Issue 2012:6)
- USAID. Training Module: How to Work with USAID Monitoring, Evaluation, and Learning
- Cassidy, Caroline and Louise Ball (2018) Communications monitoring, evaluating and learning toolkit. Overseas Development Institute

6.2 Make the message relevant and actionable

To ensure that the message is relevant and actionable, it is recommended to:

- Invite relevant stakeholders to comment on the ToC logic and integrate their feedback into the ToC and intervention design.
- Observe real-world behaviour rather than merely survey experiment results to come to solid conclusions about messaging (see Comerford, 2022).
- Use messages about the concrete impact of corruption to stimulate the audience's desire to learn more and ultimately take action.

In incentivising people to mobilise (i.e. change their behaviours), we need to be aware that we are competing for attention and that – if people are to invest time and energy – it must be to pursue an issue that is of great value to them. This means the message has to be made as relevant as possible and has to avoid triggering or reinforcing corruption fatigue.

Messages with a local framing may incentivise a positive effect. This points to highlighting, as concretely as possible, the impact of corruption on the immediate environment of citizens.

- For example, the message about taxes (among non-pessimists) had a positive effect, likely due to its connection of corruption to the individual's financial situation.
- The message about bribery in health facilities was received with relief among some users who confessed to being too poor to afford a bribe. They liked the posters and desk signs as clear elements they could point to argue against giving a bribe to facility workers.

Such an approach can be built upon by following one of the notions extrapolated from Sharot et al.'s article (7): beliefs are more likely to change when people search for the information themselves. According to that framework, highlighting what people do not know can work to trigger information seeking. In this regard, it is possible to start by pointing people to visible elements in their environment where the footprints of corruption may be felt. These could include, for example, medicines in local health facilities being out of stock, water shortages or public schools in bad condition. It can then be asked what budget had been allocated for those services and how it might have been used.

Providing access to information where people might search for the answer themselves, and a channel to denounce or communicate their findings, might prove effective. It gives people the agency to discover the causes for failing services that are important to them and adding an actionable route to do something about this.

6.3 Work with social norms

Where the evidence suggests that corruption problems are incentivised and/or exacerbated by issues of sociality, a social norms approach might be appropriate. For instance, World Values Survey data on Indonesia suggests that

while people strongly condemn corruption, they also express low levels of trust in others and express unwillingness to get involved in different types of political action and social activism. Beliefs about other people's trustworthiness and intentions may be key to incentivising action. Revealing the real injunctive social norm might be a promising approach to test.

6.4 Develop a long-term strategic approach

Some of the studies reviewed suggest approaches that might effectively elicit positive change in corruption perceptions, willingness to denounce and even propensity to bribe. However, it must be underscored that such changes are likely only of a temporary nature. Therefore, it is essential to think about messaging in terms of long-term outcomes, and to consider how to make the intervention sustainable. This can include complementing messaging with other approaches, such as:

- Changes in other factors of the choice architecture (e.g. the regulatory framework, further informal incentives)
- Community meetings
- Cultivating networks of champions
- Engaging opinion leaders
- Credible monitoring (e.g. involving a credible threat of punishment).

As addressed, only two of the seven studies tested anti-corruption messaging in the context of a real anti-bribery intervention in the field; the results from the other studies might look different if their messages were tested in the field rather than as an experiment. While the studies give us indications of the viability of anti-corruption messaging, a stronger evidence base is needed for concrete and sustainable implications. Additionally, the analysed studies tested temporary changes following a single exposure to particular messaging. Longer-term and consistent exposure to anti-corruption messages may produce different results, and thus need to be tested. Therefore, it is crucial to test more messaging in the context of real in-field interventions in various settings and longer time-frames.

7 Bibliography

See Box 1 for the papers analysed in this study.

Further sources:

Baez-Camargo, Claudia (July 2022) *Developing anti-corruption interventions addressing social norms: Lessons from a field pilot in Tanzania*. Working Paper Nr. 40. Basel Institute on Governance. Available at: https://baselgovernance.org/sites/default/files/2022-07/220726_WP-40.pdf

Bicchieri, Cristina (2017), *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York: Oxford University Press.

Comerford, David. (2022, October 25). *A Practitioner's Guide to Leveraging Behavioral Insights*. The BE Hub. Available at: <https://www.behavioraleconomics.com/leveraging-behavioral-insights-a-guide-for-practitioners/>

Falisse, Jean-Benoît and Nastassia Leszczynska (2022) Do Anti-Corruption Messages Improve Public Service Delivery? Insights from a Lab-in-the-Field Experiment in Burundi, *The Journal of Development Studies*, 58:1, 96-114.

Johnsøn, Jesper (2012) Theories of change in anti-corruption work: A tool for programme design and evaluation. U4 Issue, 2012:6. U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute. Available at: <https://www.u4.no/publications/theories-of-change-in-anti-corruption-work-a-tool-for-programme-design-and-evaluation>

Paluck, Elizabeth Levy (2009). Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda. *Journal of Personality and Social Psychology*, 96(3), 574–587.

Peiffer, Caryn and Nic Cheeseman (2023) Message misunderstood: Why raising awareness of corruption can backfire. U4 Brief, 2023:1. U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute. Available at: <https://www.u4.no/publications/message-misunderstood-why-raising-awareness-of-corruption-can-backfire>

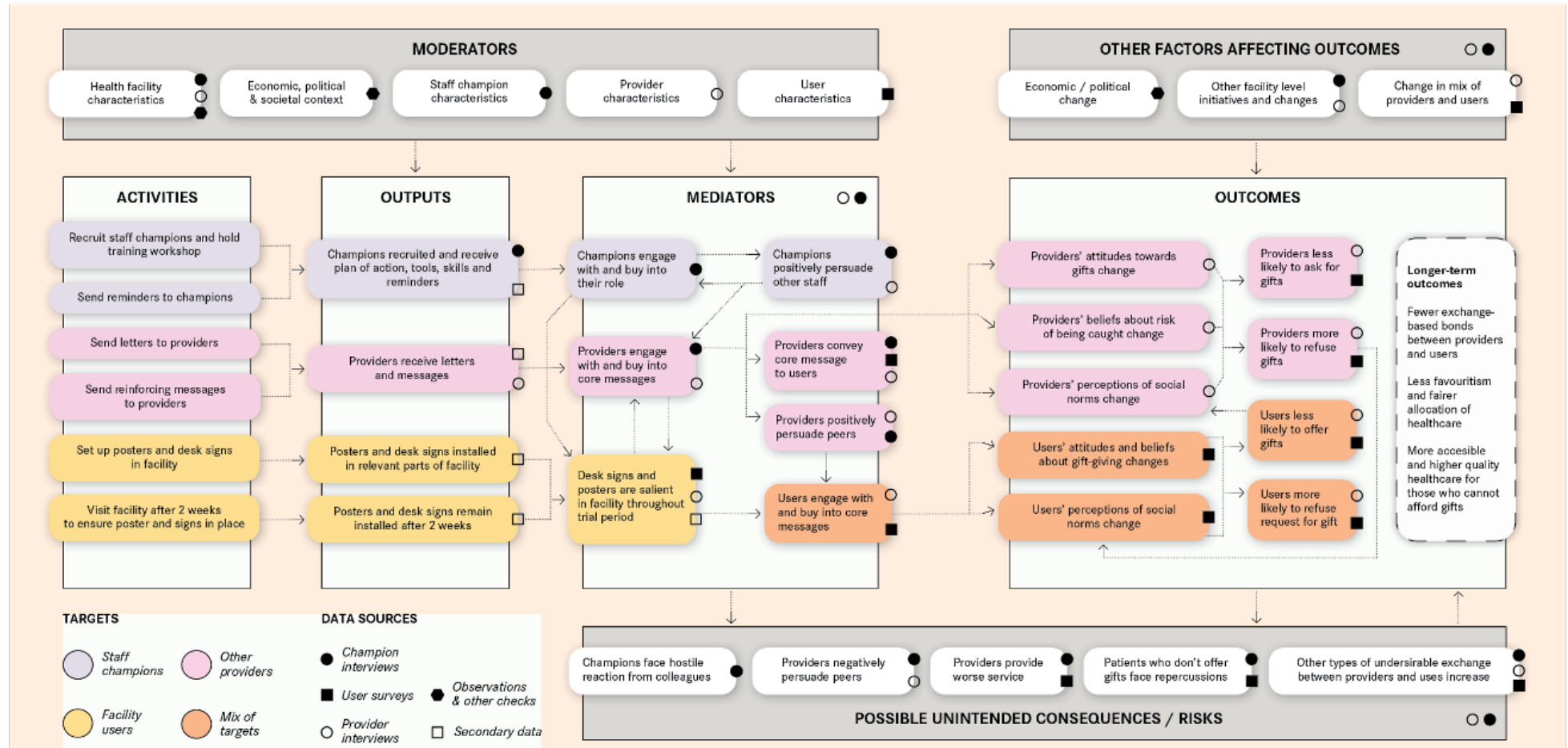
Persson, A., Rothstein, B. and Teorell, J. (2013), Why Anticorruption Reforms Fail—Systemic Corruption as a Collective Action Problem. *Governance*, 26: 449-471. <https://doi.org/10.1111/j.1468-0491.2012.01604.x>

Maclean, Catherine, John Buckell and Joachim Marti (2019) Information Source and Cigarettes: *Experimental Evidence on the Messenger Effect*. NBER Working Paper No. w25632.

Salib, Monalisa (2022) *Theory of Change Workbook: A Step-by-Step Process for Developing or Strengthening Theories of Change*. USAID. Available at: <https://usaidlearninglab.org/resources/theory-change-workbook-step-step-process-developing-or-strengthening-theories-change>

TRAFFIC and Behavioural Insights Team (2019) *Reducing Demand for Illegal Wildlife: Choosing the Right Messenger*. TRAFFIC International. Available at: <https://www.changewildlifeconsumers.org/site/assets/files/1433/dr-good-practice-guidelines-messengers-vfinal.pdf>

8 Annex



Theory of Change: Reducing gift-giving in a hospital in Tanzania. Source: Claudia-Baez-Camargo et al., 2022, p.7