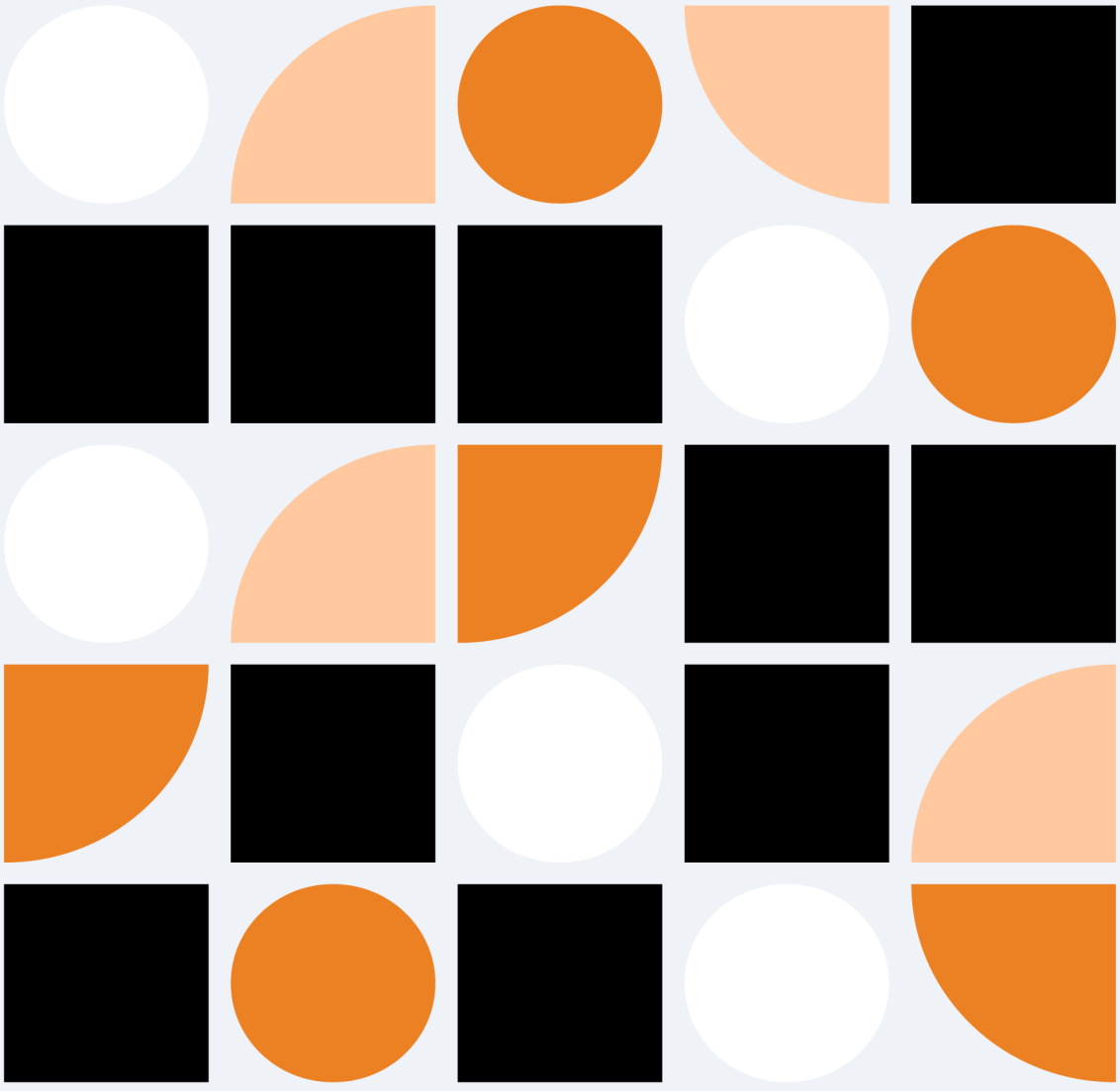
BASEL INSTITUTE ON GOVERNANCE

Behavioural insights and anti-corruption

Executive summary of a practitioner-tailored review of the latest evidence (2016–2022)

Cosimo Stahl | October 2022



Basel Institute on Governance
Steinenring 60 | 4051 Basel, Switzerland | +41 61 205 55 11
info@baselgovernance.org | www.baselgovernance.org

Associated Institute of the University of Basel

Table of contents

About the author

Cosimo Stahl

Cosimo Stahl is a Public Governance Specialist with the Basel Institute on Governance in Switzerland. He supports the Public Governance Division in developing and implementing various research and technical support projects. Email: mcjstahl@gmail.com.

About this review

This is a short version of a substantial in-depth review of the latest evidence (2016–22) on how Social Norm and Behaviour Change (SNBC) approaches can inform anti-corruption practice. This publication was supported by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH on behalf of the Federal Ministry for Economic Cooperation and Development (BMZ). The contents of this publication do not represent the official position of either BMZ or GIZ.

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Federal Ministry for Economic Cooperation and Development

Open-access licence and citation

Suggested citation: Stahl, Cosimo. 2022. “Behavioural insights and anti-corruption: Executive summary of a practitioner-tailored review of the latest evidence (2016–2022).” Basel Institute on Governance. Available at: https://baselgovernance.org/publications/behavioural-insights-and-anti-corruption.

Keywords: Anti-corruption, social norms, behaviour change, realist synthesis, behavioural insights

1 Introduction

1.1 The potential of behavioural interventions

Donors, governments and anti-corruption practitioners seeking alternative tools to address systemic corruption are increasingly turning to behavioural science. Behavioural anti-corruption approaches appear promising because they respond to a growing body of descriptive evidence on how certain social norms and mental models drive corruption, particularly in fragile contexts.[1] Interventions that target social norms and seek to shift people's behaviours away from corrupt practices could be more effective and long-lasting than ones that, for example, simply add more regulations and controls.

Yet few large-scale anti-corruption programmes have so far been informed by behavioural insights – partly due to a lack of evidence on where such an approach would be appropriate, what works and what does not.

That evidence is slowly becoming available, thanks to an increase in the past five years in what can be called Social Norms and Behaviour Change (SNBC) intervention studies. Many have yielded positive effects and demonstrate the potential of SNBC interventions to tackle systemic corruption, but some studies have encountered counterproductive effects of anti-corruption messaging.

Based on a synthesis of the evidence, this brief paper summarises a set of behavioural explanations (i.e. insights and pitfalls) for why some of these SNBC approaches have failed while others have been effective. The aim is to provide practitioners designing SNBC interventions with evidence to help them develop effective programmes and avoid common pitfalls.

The full research paper and analysis tables are available to practitioners upon request. Please email the author or info@baselgovernance.org.

[1] Cf. for anti-corruption in 'fragile' Jenkins, Martínez B. Kukutschka, and Zúñiga, 'Anti-Corruption in Fragile Settings: A Review of the Evidence'., 'adverse' Khan, Andreoni, and Pallavi, 'Anti-Corruption in Adverse Contexts'., or 'Fragile and Conflict-affected' Ventura, 'Adapting Anti-Corruption Strategies in Fragile and Conflict-Affected Settings - A Literature Review'. implementation contexts.

1.2 About this realist review

A realist review aims to assess intervention success based on the quality of a programme's theory of change and the extent to which it has been adequately contextualised. According to this logic,[2] intervention tools must be ***conceptually and empirically informed*** and tailored to the implementation context.[3] In other words, this is about assessing whether a study or intervention design adequately captures the empirical complexity surrounding concrete corrupt behaviours and whether intervention treatments have been designed accordingly. Intervention studies or actual pilot interventions should adequately and sufficiently reflect this complexity in their theories of change.[4] Expectations about how behavioural insights are expected to elicit change should be clearly specified.

1.3 The current state of the evidence (2016-22)

The literature and evidence base on behavioural approaches to anti-corruption consists of observational studies, expert guidance on operationalising SNBC approaches and experimental trials. This paper focuses on the findings concerning nine experimental SNBC studies that were shortlisted for analysis and synthesis (see

References). These SNBC studies use "behaviour change communication" (BCC) approaches in two modalities:

- The use of **information** applicable to anti-corruption public education and awareness-raising campaigns.[5]
- Harnessing of elements of **sociality,** with a shift from the individual to the collective (e.g. peer networks) as entry points to elicit change.

Some of the studies describe what might be termed "behavioural pitfalls." These are understood as weaknesses in intervention design caused by unexpected effects of

[2] Pawson et al., 'Realist Review - a New Method of Systematic Review Designed for Complex Policy Interventions'; Berg and Nanavati, 'Realist Review'.
[3] The full review report and the accompanying review protocol are available upon request.
[4] Acknowledging and factoring in contextual complexity (as in complex systems) is what behaviouralists Halpern, *Inside the Nudge Unit*; John, *How Far to Nudge? : Assessing Behavioural Public Policy*; Scharbatke-Church and Chigas, 'Using Systems Thinking to Understand and Address Corruption in the Criminal Justice System in Fragile States' for anticorruption., realist reviewers Bohni Nielsen, Lemire, and Tangsig, 'Unpacking Context in Realist Evaluations'; Pawson, *Evidence-Based Policy*; Pawson; Pawson et al., 'Realist Review - a New Method of Systematic Review Designed for Complex Policy Interventions'. and theory of change theorists (Blamey and Mackenzie 2007; Johnsøn 2012 for anti-corruption).
[5] Technically, this is about using mostly normative information heuristically (beyond merely informing), whereby exposure to such information is intended to elicit mental associations (for example, shortcutting) and/or start social learning processes.

behaviours that have not been accounted for, which directly impact the achievement of the stated intervention goals or give rise to new, undesirable outcomes.

2 Synthesis analysis

The content of each treatment (understood here as a tested approach) across the reviewed SNBC studies has been categorised according to three main review criteria:

a) Type of information employed:
   - **Normative (generic)**. [6] Such information invites people to make a value judgement about a problem such as corruption in a generic manner. The basic idea behind such use of normative information is to convey that corruption is bad but typically without adding many nuances.
   - **Normative (contextualised)**. Such information is normative, making value judgements about corruption but tailored to the context. Typically, it emphasises concrete and locally relevant problems or patterns of corruption.
   - **Socially sensitive.** Such information is sensitive and responsive to concrete social contexts.[7] It is empirically factual, addressing specific issues that matter to the target audiences and appealing to elements of sociality such as group affiliations, collective social norms and shared expectations.

b) Type of framing of the messages employed:
   - **Negative:** Information characterised by a negative tone, for example alluding to the costs of corruption or high prevailing levels of corruption.
   - **Positive:** Information that highlights success stories in the fight against corruption, that communicates hope that change is possible.

c) Breadth of the intended intervention target groups:

[6] Normative information has been referred to as summary information, say about a group [or an issue]. See: Yamin et al., 'Using Social Norms to Change Behaviour and Increase Sustainability in the Real World'. Arguably, normativity is more than that since it emphasises value judgements and normative propositions for evaluating a subject or action Jarvis Thomson, *Normativity*.
[7] Social information definable as relational, reference-dependent information relative to one's social context (based on: Hunt and Schooler, 'The Easy Part of the Hard Problem').

- **Broad:** Anti-corruption messages are destined to large, undifferentiated groups, often the entire population.
- **Narrow:** Anti-corruption messages are tailored to narrowly defined groups such as students, parents, health workers, etc.

The interventions that were considered in this review were categorised as *successful* or *unsuccessful* according to the conclusions reached by the authors. Out of the nine studies:

- ***Six*** tested messaging to be used for mass awareness raising. All these studies strike a rather negative tone in their conclusions, pointing out the potentially adverse "backfiring" effects of generic anti-corruption messaging.
- This stands in sharp contrast with the ***three*** intervention studies applying a social norms lens, all of which achieved some measure of positive change in relation to perceived norms of corruption among target groups.

2.1 Shortlisted studies (overview, by treatment)

Some interventions included more than one treatment. Across the nine intervention studies,[8] a total of 23 treatments were identified and coded against the three review criteria. Each treatment is identified by a letter code (see Table 1 on the next page).

2.2 Evidence mapping for treatment effectiveness

Different treatments applied across all studies were assessed for effectiveness. The results are categorised as be either positive **(green)**, negative **(red)**, or partial **(orange)**. All the treatments assessed are based on information testing various types of anti-corruption messages. Only one treatment also experimentally tests a peer-driven approach for the dissemination of anti-corruption messages.[9]

[8] Sometimes, there are several publications about the same intervention study. For clarity and completeness all versions have been consulted and are summarised in the Reference Table. Older (or newer) versions are marked in grey.
[9] For more detail, see: Baez-Camargo, 'Developing Anti-Corruption Interventions Addressing Social Norms - Lessons from a Field Pilot in Tanzania'; Baez Camargo et al., 'Using Behavioural Insights to Reduce Gift-Giving in a Tanzanian Public Hospital - Findings from a Mixed-Methods Evaluation'.

#	**Study** (by authors & date of publication)		**Treatment specifics**			
		Treatment code	*Target*	*Behaviour insights/pitfalls*		
				Info type	*Framing (valence)*	*Theme*
I	Denisova-Schmidt (et al. 2015)	A	narrow	Normative (generic)	Negative	Meaning, types, definitions of corruption
II	Denisova-Schmidt (et al. 2020)	B	narrow	Normative (generic)	Negative	Types & consequences of corruption
		C	narrow	Normative (generic)	Negative	Meaning, types, definitions of corruption
		D	narrow	Normative (contextualised)	Negative	Education-specific consequences of corruption
III	Corbacho (et al. 2016)	E	broad	Normative (generic)	Negative	National corruption trend
		F	broad	Normative (contextualised)	Negative	Judicial inefficiency (crime focus)
IV	Peiffer (20120, [2016])	G	broad	Normative (generic)	Negative	Grand corruption
		H	broad	Normative (contextualised)	Negative	Petty corruption (public services focus)
		I	broad	Normative (contextualised)	Positive	Governmental anticorruption success (performance/track record focus)
		J	broad	Normative (contextualised)	Neutral	Governmental call for civic action (motivational)
V	Peiffer & Walton ([2017], 2019, 2022)	K	broad	Normative (generic)	Negative	Illegality of corruption (unconstitutional)
		L	broad	Normative (contextualised)	Neutral	Religious anticorruption stance
		M	broad	Social (specific)	Neutral	Corruption as a local, community/kinship matter
		N	broad	Normative (generic)	Negative	Ubiquity of corruption
VI	Köbis (et al. 2020)	*(control)*		-	-	(service delivery bribery game)
		O	narrow	Social	Positive	regional anticorruption trend (local)
VII	Agerberg (2021)	P	broad	Social	Positive	National norm (cultural association with Mexico)
VIII	Cheeseman & Peiffer ([2020] 2021)	Q	narrow	Normative (generic)	Negative	Pervasive corruption
		R	narrow	Normative (contextualised)	Neutral	Religious anticorruption stance
		S	narrow	Normative (contextualised)	Positive	Local government anticorruption success
		T	narrow	Normative (contextualised)	Negative	Local impact of/fight against corruption
		U	narrow	Normative (contextualised)	Neutral	local tax revenue increase improving public sector performance
IX	Baez-Camargo (2022[et al. 2022)	V	narrow	Social	Positive	Positive anticorruption trend
		W championing	narrow	Social	Positive	Appeal to professional ethics/code of conduct, disseminated through network

Table 1: Overview of reviewed studies (by treatment)

Summarised in Table 2 below are all treatments, including those that worked and those that did not, categorised by type of information used, namely:

Treatment effectiveness	**Normative information** ('general')	**Normative information** ('specific')	**Social information**	*Total*
Positive effect	-	**2**	**4**	6
Partial effect	-	**1**	-	1
Negative or null effect	**13**	**3**	-	16
Total	13	6	4	**23**

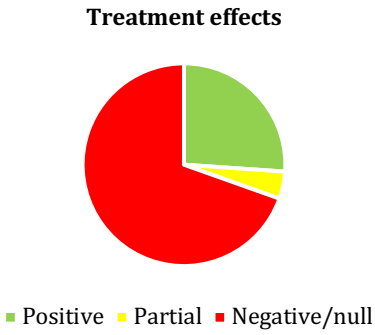


Table 2: Treatment effectiveness (by type/use of information)

In Figure 1 below, all treatments are plotted across the three effectiveness criteria.

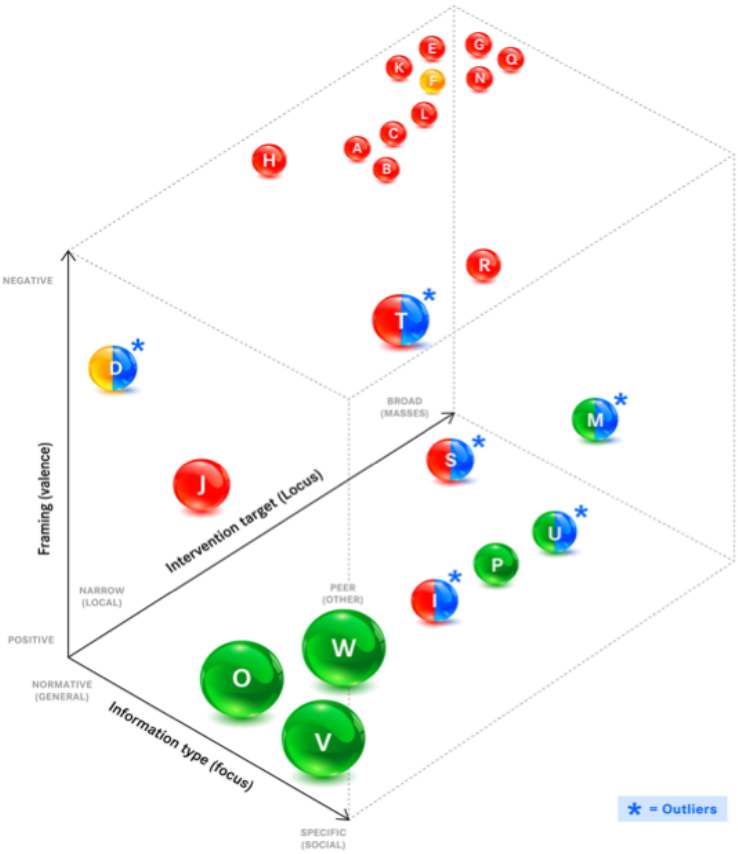


Figure 1: Visualisation of treatments against the three review criteria

This visualisation suggests that intervention treatments are more successful when they adopt **socially sensitive** messages, when they adopt more **positive framings** and when they target **narrower groups**. There were nevertheless six treatments that did not conform to this pattern, which are marked as “outliers”. These six outliers demand further attention to explore why the results deviate from the more general pattern.

3 Entry points for behavioural anti-corruption

The context and content of communication are intrinsically linked, meaning that one cannot assess the effectiveness and attributes of one without the other. In this section, the findings of the review are presented and discussed to tease out entry points for behavioural policy uptake, around:

- The use of ***behavioural insights*** to inform SNBC intervention design and to enhance treatment effects.
- The identification and mitigation of ***behavioural pitfalls*** that could render treatment effects ineffective or even counter-effective.

3.1 What does *not* work?

As Table 2 and Figure 1 above illustrate, most of the reviewed studies report null or negative results of the treatments they tested. Following the review criteria described in section 2, this section discusses the elements that appear to account for the meagre results obtained by those studies.

3.1.1 Avoid the use of generic normative information

Across intervention studies, there are 13 treatments (see Table 1 & 2 above and

References) that use normative information in a very similar fashion: they all frame various issues relative to corruption by placing emphasis on:

- The general **meaning of corruption** clarifying types (for example, petty or grand corruption) and definitions, as well as the (negative) **consequences of corruption,**[10] regionally,[11] for society as a whole,[12] or for certain demographics such as students.[13]
- The **ubiquity and pervasiveness** of corruption.[14]
- **Negative trends** of national corruption prevalence.[15]

To appeal to as broad an audience as possible, corruption is often framed in these treatments as a general issue in terms everybody can understand. In several cases this means focusing on definitions as well as the causes and consequences of corruption. Public education campaigns aim to move people to action by raising awareness about the severity of the problems of corruption. Often, this is articulated through a combination of using value-laden information about the evils of corruption and framing corruption as a widespread phenomenon.[16]

Empirical studies suggest that such approaches often backfire because, instead of moving people to action, they confirm and even exacerbate widespread beliefs about the unavoidability and normality of corruption. This may reinforce a sense of futility among target audiences, causing them to choose apathy as opposed to action. Effectively, this would mean that people's awareness about corruption is indeed raised, but adversely so in the direction of fatalism and inaction.

3.1.2 Abstain from reinforcing *negativity* when framing corruption

Another complementary explanation has to do with the tone of anti-corruption messaging, or *information valence*. Definable as the *affective intrinsic quality of an event, object, or situation,*[17] the tone of corruption messaging is by default negative.

[10] Treatments A, B, C, G, H, K, L.
[11] Treatments Q and U.
[12] Treatments G and N.
[13] Treatments A and C.
[14] Treatments N and Q.
[15] Treatments E, F, K, and L.
[16] Issue-framing is otherwise used by the media or in politics for example Badie, Berg-Schlosser, and Morlino, 'Policy Framing'; Entman, 'Framing'; Iyengar, 'Framing Responsibility for Political Issues'., raising the salience of the normative elements surrounding a public or policy issue.
[17] See Frijda (1986), for instance.

Table 2 below shows whether messaging has been framed in a positive, negative or neutral tone in each of the reviewed treatments. The fact that not one of the negatively framed information treatments has been effective speaks for itself.

Treatment effectiveness	**Negative framing** (tone)	**Neutral framing** (tone)	**Positive framing** (tone)	*Total*
Positive effect	-	**2**	**4**	**6**
Partial effect	**1**	**-**	-	**1**
Negative or null effect	**10**	**4**	**2**	**16**
Total	11	6	6	**23**

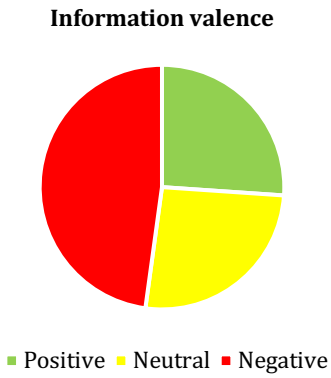


Table 1: Treatment effectiveness (by valence)

The takeaway lesson might be that emphasising overly negative messages leads to the term corruption becoming so laden with negative mental associations that it not only cuts short the desired persuasion or learning effects but lends itself to partisan exploitation and manipulation. Other areas of corruption-specific communication corroborate this.[18]

For the sake of influencing public opinion, this has forced communicators to find alternative positive formulations involving, for example, integrity and ethics.[19] It is not clear, however, that more positive formulations consistently work better. In fact, one of the reviewed treatments (outlier I) shows that corruption messaging can still backfire even if it is positively framed. This case, from Indonesia, stressed the successful track record of the national anti-corruption agency (KPK).

[18] This point of concern has been raised by the literature on political contestation and electoral framing around the issue of corruption Curini and Martelli, 'A Case of Valence Competition in Elections Parties' Emphasis on Corruption in Electoral Manifestos'; Green, 'When Voters and Parties Agree'; Stokes, 'Spatial Models of Party Competition'; also Stahl, 'Claiming and Framing for Anticorruption? Assessing the Mobilising Effects of Brazil's Public Federal Ministry's Communication Campaign during the Lava Jato Investigations (2014-21)'.
[19] For example, Badie, Berg-Schlosser, and Morlino, 'Policy Framing'; Byrne, Arnold, and Nagano, 'Building Public Support for Anti-Corruption Efforts'.

3.1.3 Do not neglect articulating detailed and empirically informed theories of change

The reviewed treatments also suggest that more general and vaguely defined intervention goals and targets are associated with unsuccessful outcomes. This suggest the importance of narrowly defining the scope of an intervention as well as explicitly thinking through how any change is expected to come about. Proper thought must be given to the role of context in shaping both corrupt behaviours *and* the assumptions about behaviour change that SNBC anti-corruption interventions envision. As a general point of advice: the choice of intervention content must resonate with target audiences.

Most of the reviewed studies appear not to have been precise in fleshing out detailed theories of change, other than in specifying, in their respective theoretical frameworks, the nature of the corruption problem at hand.[20] Theories of change are needed because they help to spell out and test expected pathways to eliciting change in the desired behaviours.[21] By doing so, theories of change enable learning about what has worked where and why. They also help us to understand what expectations did not work out and why.

Theories of change for SNBC interventions should be informed by findings stemming from relevant fields of study. These include prominently behavioural economics and clinical psychology, but other disciplines might provide useful frameworks as well. For example, communication theories and research on framing suggest that effective use of information for any act of public communication is about developing the "right" *content* relative to *context.*[22] Moreover, communicators should give due thought to the quality and extent of *consistency, credibility,* and *resonance* of the information that is to be communicated.[23]

[20] Most intervention studies conceptualise problems of mostly systemic and pervasive corruption, rather than specify how change is brought about (with few exceptions) or targeting narrowly defined behaviours as opposed to "corruption" as an umbrella term.
[21] One should think about these various pathways for each and every treatment. For example, check out: Baez Camargo et al., 'Using Behavioural Insights to Reduce Gift-Giving in a Tanzanian Public Hospital - Findings from a Mixed-Methods Evaluation', 7.
[22] Cf. Stahl (forthcoming), 'Claiming and Framing for Anticorruption? Assessing the Mobilising Effects of Brazil's Public Federal Ministry's Communication Campaign during the Lava Jato Investigations (2014-21)'.
[23] For example, is the content of communication consistent over time (i.e. frequency)? Also, is such communication consistent with the communicators' actions, mandate and competency? As such, does its content resonate both empirically (with what is happening) and publicly (with what public audiences care about)? Lastly, are both communication and communicator credible? The latter is thus not only a question of content but also a matter of *who does the* communicating – as well as *how*, *when* and *where*. These so-called three principles of communication likely interact and should therefore be considered conjointly by communication campaigners and strategists (ibid.).

3.2 What *does* work?

This section is dedicated to distilling lessons about what appear to have been effective uses of behavioural insights to inform anti-corruption programming.[24] It also suggests ways in which behavioural elements might be effectively applied, summarised in the form of four promising entry points for advancing behavioural programming.

3.2.1 Preferably use *social* information for addressing social norms

Overall, treatments that aimed to update social norms about corruption were effective. Essentially, most of these treatments moved people into thinking more positively that attitudinal and behaviour change vis-a-vis corruption is possible. In this regard, these studies indicate that, to instil and reinforce a sense of collective change, the use of information that is relevant relative to concrete social contexts proves crucial. Such information should be empirically factual and address specific issues that resonate with target audiences.

To enhance the effectiveness of treatments dealing with social norms it is key to make the information disseminated as credible as possible. As this review suggests, advertising a positive trend – such as a regional decrease in bribery – can involve citing a trusted source (such as Transparency International), for instance, or certain public institutions (a renowned university) to boost the credibility of the claim. This may prove vital in low-trust, low-cooperation contexts.

3.2.2 Make normative information relevant to the context

Normative messaging can work if it is tailored to specific target audiences. This involves framing messages about corruption in such a way that they resonate with a specific demographic (such as students) or a local target audience (such as a community). For example, one can frame issues of corruption in the education sector, and by so doing effectively tailor messages to a narrow intervention audience affected by a particular pattern of corruption. Notably, such a campaign does not target the whole population and is narrower in scope.

[24] What has effectively worked both in terms of actual intervention studies and the individual treatments across studies is graphically depicted in the main report, which is available upon request.

Alternatively, one can frame local forms of corruption as a community issue, linking pieces of normative information about corruption to problems or issues that citizens care about.

Lastly, one can also try to anchor normative information by drawing from positive developments in context, such as a regional increase in tax expenditure (at the local government level) that palpably boosted public-sector performance, for the benefit of local communities.

3.2.3 Narrowly identify intervention targets

Not all patterns of corruption are equally relevant or feasible to tackle. Therefore, the focus of any anti-corruption intervention *and* its treatments should be strategically chosen and should be narrow, zooming into specific modalities of corruption that happen regularly in repeated patterns. This also includes finding the right intervention locus – meaning where and with whom the intervention takes place – which adequately reflects where and how corruption happens. [25]

The types of corruption targeted as well as the outcomes expected from the intervention should ideally also be relevant to their respective contexts. Such choices should stem from an informed operational judgement, clearly articulated in the theory of change of the programme.

A narrow focus on actual norms and/or practices of corruption that intervention audiences are known to experience and care about should go hand in hand with a communal, cultural, and/or sectoral intervention locus:

- By targeting kinship communities locally, as was done in provincial South Africa, where descriptive norms of corruption – the widespread perception of one's social environment as highly corrupt – drive bribery in the public service sector directly. [26]
- By specifically appealing to a strong sense of cultural community among a target reference group, as was done among Mexicans for commonly labouring under widespread misperceptions of corruption as the injunctive norm. Messaging that

[25] Based on Heywood (2017), who defines the *locus* as the primary unit of analysis indicating the level, direction and location of corruption, whereas the *focus* specifies the type, modality or even sector of corruption (p. 21, 47). The meaningful choice of focus and locus, too, go hand in hand.
[26] Treatment O.

factually corrects the misperception among fellow Mexicans that corruption is an inherent part of national Mexican culture has proven effective.[27]

- By addressing a specific demographic of service seekers and providers for experiencing social norms of reciprocity that incentivises gift-giving, a common practice turned bribery proxy, in Tanzania's health sector.[28]

Thus, the focus on a narrow target behaviour can in itself contribute to a sense of localisation among target audiences. For example:

- The South Africa study above localises bribery for a locally prevailing practice that is frequent but to a palpably lesser degree, which is a trend specific to KwaZulu-Natal. Furthermore, it targets the tight-knit urban community of Manguzi to ensure community involvement and dissemination.[29]
- The Tanzania pilot intervention targets the proactive use of bribes in network settings, notably by health facility users to befriend providers and receive better treatment, at a top Dar es Salaam hospitals that is much know to and visited by local communities.[30]

These two latter studies tailored their intervention design to target narrowly defined groups of public officials and citizens, as well as providers and users of health services respectively. This narrow targeting likely contributes to resonance among intervention audiences. So does focussing on a sector such as public health and services for that they are crucial in improving and sustaining people's livelihoods.[31]

In sum, this review makes a case for a) narrowly addressing concrete corrupt behaviours that are clearly targeted in the exact locations where they occur; b) making normative messaging locally relevant, again narrowly identifying the type of corruption that is being targeted. Both approaches aim to make programming resonate with the intended target groups and beneficiaries of the intervention.

[27] Arguably, this cultural cue may also have a localising affect, for instance, as a potential appeal to national pride about Mexican culture.
[28] Treatment V and W.
[29] Citizen-official interactions at the point of service delivery are game-theoretically modelled, and then role-played by randomly assigned participants taking up the role of either.
[30] In Baez-Camargo's study, real individuals pertaining to either group are selected for treatment.
[31] On the sustainable livelihood strategies, see: DFID, 'Sustainable Livelihoods Guidance Sheets'; or Gilling, Jones, and Duncan, 'Sector Approaches, Sustainable Livelihoods and Rural Poverty Reduction'.

3.3 The importance of *nuanced* contextualisation

The question remains whether intervention studies have meaningfully factored contextual elements into their intervention or study design and whether this may have been consequential for treatment outcomes and effectiveness.

3.3.1 Degrees and qualities of contextualisation

Most of the newer studies, identified below by their letter codes as noted in the

References, explicitly draw on key structural and situational aspects of local contexts and circumstances. They single out and build upon specific context characteristics, notably at two levels:

- At the ***structural*** level, considering the influence of institutional, social or cultural (to a lesser degree) *orders* and *systems* upon patterns of pervasive corruption. For example:
    - The alleged cultural appropriation of corruption as the norm.[32]
    - Gross inefficiency of judicial institutions.[33]
    - A political zero tolerance stance on corruption and resulting fear of reprisal.[34]
    - Community systems such as *Wantok*.[35]
- At the more ***situational*** level, with regards to specific circumstances, such as events, developments, or trends. This includes:
    - Governmental success in fighting corruption.[36]
    - Improvement in local government performance.[37]
    - *Positive* regional anti-corruption trends.[38]
    - *Negative* trend of political corruption.[39]
- At the **individual** level in terms of targeting specific groups and their presumed individual ***predispositions***:
    - Law-abiding propensities

[32] Treatment P.
[33] Treatment F.
[34] Treatments V and W.
[35] Treatment M.
[36] Treatments J and T.
[37] Treatments I and U.
[38] Treatments O.
[39] Treatment E.

        - Moral, righteous and legal-positivist propensities (illegality/crime-adverse, law-abiding).[40]
    - Moral, ethical codes
        - From strong religious (anti)corruption beliefs/views.[41]
    - Local community pressures and expectations:
        - With expressions such as Wantok, indicating degree of social proximity and group belonging, and thus a series of implicit obligations.[42]
    - Norms incongruence
        - From collectively misperceiving what others do, and/or think is right and wrong.[43]
    - Social expectations (service seekers) from peer pressure and strong professional ethics (service providers/medical staff).
        - From a sense of professionalism and professional identity.[44]

While these are distinct analytical categories, they often interact on the ground. Accounting for their contextual compounding effects in theory of programme and change thinking can prove crucial.

3.3.2 Invest in a *nuanced* understanding of context

From an SNBC perspective, proper contextualisation is about capturing the nuances in the contexts where relevant problems of corruption are prevalent across three spheres:

1) **Macro-level structural determinants of routine behaviours.** [45] These include elements such as weak state capacity, political instability, socio-economic hardship, for driving and normalising economic improvisation and social bricolage, [46] for instance.

[40] Treatment K.
[41] Treatments L and R.
[42] Treatment M, for example.
[43] Treatment P.
[44] Treatment W.
[45] These behaviours tend to become engrained on a larger scale, by means of routinisation and normalisation through extrinsic experience and repeated interaction with other members of the larger society. For example, haggling is a typical routine behaviour from a repertoire of learned and enculturated economic improvisation practices. As such, it may be a more stable and more change resistant behaviour.
[46] Such structural deficiencies and systemic weaknesses force most people to largely resort to informal means on an organisational basis, both in economic terms (through an informal economy of favours, for example) and in social terms (for instance, by building informal social networks to get by).

2) **Meso-level social elements in the immediate choice environment.**[47] These include elements such as social norms, peer pressure, shared stereotypes.
3) **Individual-level cognitive patterns.** These include elements such as cognitive biases, automatic thinking as well as personal attitudes and preferences.[48]

In practice the three are likely to be interdependent and hard to measure with exactitude. However, it is important to keep them in mind as distinct categories in the process of developing theories of change because they provide entry points to specific intervention elements. Often issues at all levels will need to be addressed in order to tackle persistent problems of corruption. For example, bribery might arise out of structural factors such as under-resourced public services, but be reinforced through social norms of reciprocity and individually held double standards ("when others do it is corruption, but in my case it is morally justified"). Or else social distrust and repeated first-hand experience of corruption may produce *collective apathy* and *cynicism*, causing people to turn into "pessimist believers".[49]

We must assume that populations selected for treatment are susceptible to a combination of these contextual variables. Addressing these programmatically would require targeted approaches that build on nuanced context analysis, according to leading behavioural scientists.[50]

4 Concluding remarks

This review synthesis has aimed to identify what lessons we can learn from selected SNBC anti-corruption interventions. The main message is that anti-corruption approaches need to be **narrowly defined** in scope, targeting **concrete patterns** of corruption that are **relevant to the context.** A thorough intervention design must specify the **pathways** through which the intervention elements are expected to elicit a change in the targeted behaviours. Only by

[47] This concerns more immediate choice and decision-making, which is often situationally specific. Such choices are subject to change on an ad hoc basis.
[48] These tend not to hinge much upon context and environment; instead, they are often intrinsic to a person's motivation and personality.
[49] See: Cheeseman and Peiffer, 'The Curse of Good Intentions'..
[50] Halpern, *Inside the Nudge Unit*; John, *How Far to Nudge? Assessing Behavioural Public Policy*.

empirically testing whether those expected pathways of change are validated in practice can we build robust evidence on what works to address entrenched patterns of corruption.

Delving into the evidence characterising the context and accounting for what the literature tells us about behavioural mechanisms and possible behavioural pitfalls should be elements informing SNBC interventions. There is still much to learn from other areas of policymaking, behavioural science and beyond. Therefore, encouraging interdisciplinarity in the testing of empirical approaches can be helpful to develop evidence-based models that can shed light on how systems can transit from high-corruption to low-corruption equilibria.

As more anti-corruption SNBC intervention approaches are tested, practitioners are encouraged to transparently communicate their results as well as detailed descriptions of their methodologies. Only by sharing experiences, successful or not, can the evidence base to inform more effective programming decisions be expanded and important lessons learned

5 References

5.1 Shortlisted and reviewed publications

#	Study (by authors & date of publication)	Bibliographical information	
		Citation	***DOI/Link***
I	Denisova-Schmidt (et al. 2015)	Denisova-Schmidt, E., Huber, M., Prytula, Y., 2015. An experimental evaluation of an anti-corruption intervention among Ukrainian university students. *Eurasian Geography and Economics, 56*: 6, 713-734.	https://doi.org/10.1080/15387216.2016.1155467
II	Denisova-Schmidt (et al. 2020)	Denisova-Schmidt, E., Huber, M., Leontyeva, E., Solovyeva, A., 2020. Combining experimental evidence with machine learning to assess anti-corruption educational campaigns among Russian university students. *Empirical Economics 60*, 1661–1684.	https://doi.org/10.1007/s00181-020-01827-1
		Denisova-Schmidt, E., Huber, M., Leontyeva, E., 2016a. On the development of students' attitudes towards corruption and cheating in Russian universities. *European Journal of Higher Education 6: 2*, 128–143.	https://doi.org/10.1080/21568235.2016.1154477
III	Corbacho (et al. 2016)	Corbacho, A., Gingerich, D.W., 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*, 1077–1092.	https://doi.org/10.1111/ajps.12244
IV	Peiffer (2020, [2016])	Peiffer, C., 2020. Message Received? Experimental Findings on How Messages about Corruption Shape Perceptions. British Journal of Political Science 50, 1207–1215.	https://doi.org/10.1017/S0007123418000108
		Peiffer, C., 2016. *Message Received: how messages about corruption shape perceptions*. Paper prepared for 2016 Development Studies Association (DSA) conference ('Politics in Development'). 12-14 September. Oxford (UK).	https://nomadit.co.uk/conference/dsa2016/paper/33619/paper-download.pdf
V	Peiffer & Walton ([2017], 2019, 2022)	Peiffer, C., Walton, G., 2022. Getting the (right) message across: How to encourage citizens to report corruption. *Development Policy Review 40: 5, e12621.*	https://doi.org/10.1111/dpr.12621
		Peiffer, C., Walton, G., 2019. Overcoming Collective Action Problems through Anti-Corruption Messages. *Development Policy Centre Discussion Paper no 77*. Available at SSRN.	https://doi.org/10.2139/ssrn.3333475
		Walton, G.W., Peiffer, C., 2017. The impacts of education and institutional trust on citizens' willingness to report corruption: lessons from Papua New Guinea. *Australian Journal of Political Science 52: 4*, 517–536.	https://doi.org/10.1080/10361146.2017.1374346
VI	Köbis (et al. 2019)	Köbis, N.C., Troost, M., Brandt, C.O., Soraperra, I., 2019. Social norms of corruption in the field: social nudges on posters can help to reduce bribery. *Behavioural Public Policy 6:4*, 1–28.	https://doi.org/10.1017/bpp.2019.37
VII	Agerberg (2021)	Agerberg, M., 2021. Messaging about corruption: The power of social norms. *Governance 35:3,* 929-950.	https://doi.org/10.1111/gove.12633
VIII	Cheeseman & Peiffer ([2020] 2021)	Cheeseman, N., Peiffer, C., 2021. The Curse of Good Intentions: Why Anticorruption Messaging Can Encourage Bribery. *American Political Science Review 116:3*, 1081-1095.	https://doi.org/10.1017/S0003055421001398
		Cheeseman, N., Peiffer, C., 2020. *The unintended consequences of anti-corruption messaging in Nigeria: Why pessimists are always disappointed.* ACE - Anti-Corruption Evidence (Working Paper). SOAS - University of London, London.	https://ace.soas.ac.uk/publication/the-unintended-consequences-of-anti%E2%80%90corruption-messaging-in-nigeria-why-pessimists-are-always-disappointed/
IX	Baez-Camargo (2022, [et al. 2022])	Baez-Camargo, C. 2022. *Developing anti-corruption interventions addressing social norms: Lessons from a field pilot in Tanzania.* Working Paper 40, Basel Institute on Governance. University of Basel.	https://baselgovernance.org/publications/wp-40
		Baez Camargo, C., Gadenne, V., Mkoji, V., Perera, D., Persian, R., Stark, 2022. *Using behavioural insights to reduce gift-giving in a Tanzanian public hospital – Findings from a mixed-methods evaluation.* Technical Research Report. Basel Institute on Governance, University of Basel.	https://baselgovernance.org/publications/TZ-giftgiving

5.2 Secondary sources

Badie, Bertrand, Dirk Berg-Schlosser, and Leonardo Morlino. 'Policy Framing'. In *International Encyclopedia of Political Science*, 1924–27. Thousand Oaks: SAGE Publications, Inc., 2011. https://doi.org/10.4135/9781412994163.

Baez Camargo, Claudia, Violette Gadenne, Veronica Mkoji, Dilhan Perera, Ruth Persian, and Stark. 'Using Behavioural Insights to Reduce Gift-Giving in a Tanzanian Public Hospital - Findings from a Mixed-Methods Evaluation'. Basel Institute on Governance - University of Basel, September 2022. https://baselgovernance.org/publications/TZ-giftgiving.

Baez-Camargo, Claudia. 'Developing Anti-Corruption Interventions Addressing Social Norms - Lessons from a Field Pilot in Tanzania'. Basel Institute on Governance - University of Basel, May 2022.

Berg, Rigmor C., and Julie Nanavati. 'Realist Review: Current Practice and Future Prospects'. *Journal of Research Practice* 12, no. 1 (8 July 2016): R1–R1.

Blamey, Avril, and Mhairi Mackenzie. 'Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges?' *Evaluation* 13, no. 4 (1 October 2007): 439–55. https://doi.org/10.1177/1356389007082129.

Bohni Nielsen, Steffen, Sebastian Lemire, and Stinne Tangsig. 'Unpacking Context in Realist Evaluations: Findings from a Comprehensive Review'. *Evaluation*, 25 October 2021, 135638902110530. https://doi.org/10.1177/13563890211053032.

Byrne, Elaine, Anne-Katrin Arnold, and Fumiko Nagano. 'Building Public Support for Anti-Corruption Efforts : Why Anti-Corruption Agencies Need to Communicate and How'. The World Bank, 1 January 2010. http://documents.worldbank.org/curated/en/2010/01/12204776/building-public-support-anti-corruption-efforts-anti-corruption-agencies-need-communicate.

Cheeseman, Nic, and Caryn Peiffer. 'The Curse of Good Intentions: Why Anticorruption Messaging Can Encourage Bribery'. *American Political Science Review*, 20 December 2021, 1–15. https://doi.org/10.1017/S0003055421001398.

Curini, Luigi, and Paolo Martelli. 'A Case of Valence Competition in Elections Parties' Emphasis on Corruption in Electoral Manifestos'. *Party Politics* 21, no. 5 (1 September 2015): 686–98. https://doi.org/10.1177/1354068813491540.

DFID. 'Sustainable Livelihoods Guidance Sheets'. London: UK Department for Int'l Development, 1999.

Entman, Robert M. 'Framing: Toward Clarification of a Fractured Paradigm'. *Journal of Communication* 43, no. 4 (1993): 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x.

Frijda, Nico H. *The Emotions*. Studies in Emotion and Social Interaction. Cambridge [etc: Cambridge University Press, 1986.

Gilling, Jim, Stephen Jones, and Alex Duncan. 'Sector Approaches, Sustainable Livelihoods and Rural Poverty Reduction'. *Development Policy Review* 19, no. 3 (2001): 303–19. https://doi.org/10.1111/1467-7679.00136.

Green, Jane. 'When Voters and Parties Agree: Valence Issues and Party Competition'. *Political Studies* 55, no. 3 (2007): 629–55. https://doi.org/10.1111/j.1467-9248.2007.00671.x.

Halpern, David. *Inside the Nudge Unit: How Small Changes Can Make a Big Difference*. Random House, 2015.

Heywood, Paul M. 'Rethinking Corruption: Hocus-Pocus, Locus and Focus'. *The Slavonic and East European Review* 95, no. 1 (2017): 21–48. https://doi.org/10.5699/slaveasteurorev2.95.1.0021.

Hunt, Tam, and Jonathan W. Schooler. 'The Easy Part of the Hard Problem: A Resonance Theory of Consciousness'. *Frontiers in Human Neuroscience* 13 (2019). https://www.frontiersin.org/article/10.3389/fnhum.2019.00378.

Iyengar, Shanto. 'Framing Responsibility for Political Issues'. *The Annals of the American Academy of Political and Social Science* 546 (1996): 59–70.

Jarvis Thomson, Judith. *Normativity*. Chicago, Ill.: Open Court, 2008.

Jenkins, Matthew, Roberto Martínez B. Kukutschka, and Nieves Zúñiga. 'Anti-Corruption in Fragile Settings: A Review of the Evidence'. GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit), September 2020. https://www.giz.de/de/downloads/giz2020_en_anti-corruption_in_fragile_states.pdf.

John, Peter. *How Far to Nudge? : Assessing Behavioural Public Policy*. Cheltenham, UK : Edward Elgar Publishing, 2018.

Johnsøn, Jesper. 'Theories of Change in Anti-Corruption Work: A Tool for Programme Design and Evaluation'. Chr. Michelsen Institute (U4), 2012. http://www.u4.no/publications/theories-of-change-in-anti-corruption-work-a-tool-for-programme-design-and-evaluation/.

Khan, Mushtaq, Antonio Andreoni, and Roy Pallavi. 'Anti-Corruption in Adverse Contexts: Strategies for Improving Implementation'. Working Paper. ACE Anti-Corruption Evidence (SOAS Consortium). London: SOAS, September 2019. https://ace.soas.ac.uk/publication/anti-corruption-in-adverse-contexts-strategies-for-improving-implementation/.

Pawson, Ray. *Evidence-Based Policy: A Realist Perspective / Ray Pawson*. London: SAGE, 2006. https://methods.sagepub.com/book/evidence-based-policy.

Pawson, Ray, Trisha Greenhalgh, Gill Harvey, and Kieran Walshe. 'Realist Review - a New Method of Systematic Review Designed for Complex Policy Interventions'. *Journal of Health Services Research & Policy* 10, no. 1_suppl (1 July 2005): 21–34. https://doi.org/10.1258/1355819054308530.

Scharbatke-Church, Cheyanne, and Diana V. Chigas. 'Using Systems Thinking to Understand and Address Corruption in the Criminal Justice System in Fragile States'. In *Corruption, Social Sciences and the Law*. Routledge, 2019.

Stahl, Cosimo. 'Claiming and Framing for Anticorruption? Assessing the Mobilising Effects of Brazil's Public Federal Ministry's Communication Campaign during the Lava Jato Investigations (2014-21)'. PhD Dissertation, University College London, forthcoming.

Stokes, Donald E. 'Spatial Models of Party Competition'. *American Political Science Review* 57, no. 02 (June 1963): 368–77. https://doi.org/10.2307/1952828.

Ventura, Rosemary. 'Adapting Anti-Corruption Strategies in Fragile and Conflict-Affected Settings - A Literature Review'. (CJL) The Fletcher School - Tufts University, December 2021. https://www.corruptionjusticeandlegitimacy.org/items/adapting-anti-corruption-strategies-in-fragile-and-conflict-affected-settings%3A-a-literature-review.

Yamin, Paulius, Maria Fei, Saadi Lahlou, and Sara Levy. 'Using Social Norms to Change Behavior and Increase Sustainability in the Real World: A Systematic Review of the Literature'. *Sustainability* 11, no. 20 (January 2019): 5847. https://doi.org/10.3390/su11205847.