



U4 Issue 2022:6

Effectively evaluating anti-corruption interventions

Tailoring the approach to the challenge

Author
Cecilie Wathne

Series editor
Peter J. Evans

CMI CHR.
MICHELSEN
INSTITUTE

Disclaimer

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

Partner agencies

German Corporation for International Cooperation – GIZ

German Federal Ministry for Economic Cooperation and Development – BMZ

Global Affairs Canada

Ministry for Foreign Affairs of Finland

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

Swedish International Development Cooperation Agency – Sida

Swiss Agency for Development and Cooperation – SDC

The Norwegian Agency for Development Cooperation – Norad

UK Aid – Foreign, Commonwealth & Development Office

About U4

U4 is a team of anti-corruption advisers working to share research and evidence to help international development actors get sustainable results. The work involves dialogue, publications, online training, workshops, helpdesk, and innovation. U4 is a permanent centre at the Chr. Michelsen Institute (CMI) in Norway. CMI is a non-profit, multi-disciplinary research institute with social scientists specialising in development studies.

www.U4.no

U4@cmi.no

Cover photo

iStock.com/ppi09 (CC pd) <https://www.istockphoto.com/photo/50-euro-gm482840377-37279188>

Keywords

anti-corruption measures - anti-corruption strategy - anti-corruption tools - impact evaluation

Publication type

U4 Issue

Creative commons

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

Evaluating anti-corruption interventions is not easy. Challenges include how to accurately measure corruption and changes in corruption; prove causation and contribution; and to fairly and fully gauge results and sustainability. Combating corruption is a complex, long-term undertaking. The possibility of unintended consequences and backlashes is ever present. Evaluators should ensure their methodology is designed to overcome theme-specific challenges.

Main points

- Corruption is a complex and clandestine phenomenon, sustained by underlying drivers, entrenched interests, and power relations. Subsequently, evaluating anti-corruption interventions presents a myriad of challenges.
- Evaluation methodology must consider the difficulty of accurately measuring corruption and changes in corruption; stakeholders' potential reluctance to discuss the topic honestly and openly; the problem of proving causation and contribution; the time taken to achieve an impact; and the likelihood of unintended consequences and backlashes.
- Conventional anti-corruption approaches have not consistently yielded anticipated results. Assessment of effectiveness, impact, and sustainability should therefore include whether the intervention aligns with the latest and local thinking on corruption and anti-corruption measures.
- Evaluations can contribute to filling the anti-corruption evidence gap. It is possible to carry out a high-quality impact evaluation, even with budget and data limitations, by appreciating the range of methods available – including those particularly appropriate for complex interventions.
- Evaluations that focus on impact and sustainability should gauge whether the anti-corruption intervention contributes to wider, and deeper, change. With widespread or systemic corruption, eliminating a specific corrupt practice is not enough. The root causes of corruption need to be addressed in order to bring about systems change.

Table of contents

Choosing the best approach to ensure effective evaluation	1
Part 1: Understanding corruption and anti-corruption efforts	1
1.1 There are many forms of corruption – the scale and types of corruption, as well as the actors involved and the level of harm done, vary by context.	2
1.2 There are many anti-corruption tools, the details of which matter	2
1.3 A unique combination of approaches, tools, and actors is needed to address the (often) multiple reinforcing drivers of corruption in a given context	3
1.4 When corruption is systemic, anti-corruption efforts need to take a systems approach	4
1.5 Anti-corruption interventions need to be flexible, politically responsive, and designed with potential backlashes in mind	6
1.6 There are limits to what anti-corruption interventions alone can achieve, as well as the role that donor agencies can play, so successful anti-corruption efforts may require a broader approach	8
Part 2: Implications for anti-corruption evaluations	9
2.1 Consider the need for a complexity-responsive evaluation methodology	11
2.2 Carefully gather and make use of data	14
2.3 Employ a realistic barometer of intervention success – even if this means questioning the logframe	18
2.4 When attribution cannot be determined, aim to demonstrate contribution	19
2.5 Assess the extent to which there is a problem–strategy match	20
2.6 Determine whether the intervention is in line with latest and local thinking on corruption and anti-corruption efforts	21
2.7 Explore potential unintended consequences and the likelihood of a backlash	22
2.8 Draw on all of these insights, and more, to assess sustainability	23
2.9 Determine impact and ‘what works’, using appropriate anti-corruption methods	24
2.10 Assess effectiveness and impact at the appropriate level	29
Filling the anti-corruption evidence gap – the benefits of effective evaluations	30
References	32

About the author

Cecilie Wathne

Cecilie Wathne led U4's thematic work on Measurement and Evaluation from 2018 to 2020. She has been the head of planning, monitoring, evaluation, accountability and learning at the Norwegian NGO Strømme Foundation. She has also worked as a researcher at the Overseas Development Institute on issues related to aid effectiveness and accountability. Wathne holds an MSc in Economics for Development from the University of Oxford and a BSc in Economics, a BA in Political Science and a Certificate in Quantitative Methods from the University of Washington. She has undertaken fieldwork in over a dozen countries in Asia and Africa.

Acknowledgements

The author would like to thank Mari Norbakk and Maria Helena Meyer Dolve for their background research, as well as the reviewers for their expertise and advice.

Choosing the best approach to ensure effective evaluation

Corruption is complex, entrenched, and clandestine. Evaluating anti-corruption interventions is therefore not easy. Key challenges include the difficulty of accurately measuring corruption and changes in corruption; the time it takes to achieve an impact; the difficulty of proving causation and contribution; and the likelihood of unintended consequences and backlashes.

This Issue aims to provide theme-specific advice for anyone commissioning or carrying out an evaluation that either exclusively or in part assesses anti-corruption interventions. The target audience is individuals familiar with evaluations but with limited expertise in corruption and anti-corruption measures. It may also prove insightful for anti-corruption experts with limited evaluation knowledge. The focus is on interventions that aim to reduce corruption in society – not corruption risk management – and, in particular, those that are donor-funded. However, most insights are generally applicable.

The paper is intended as a complement to evaluation handbooks. The headline messages will be obvious to most readers. Carefully gathering and making use of data is, for example, something every evaluator knows how to do. However, they may not know the specific data challenges with regard to corruption and anti-corruption initiatives – and how they can be overcome. Part One provides an introduction to corruption and anti-corruption measures; Part Two explores the implications for evaluations.

Since it is not intended as a stand-alone handbook, the Issue does not cover all of the messages emphasised in comprehensive evaluation handbooks. For a more overarching introduction to evaluations, see Sida's 2020 [Evaluation Handbook](#) or UNODC's 2017 [Evaluation Handbook](#).

Part 1: Understanding corruption and anti-corruption efforts

In order to appreciate the theme-specific implications for evaluations, we must first understand the field. The [U4 Issue *Understanding corruption and how to curb it: A synthesis of latest thinking*](#) identifies ten main insights into

corruption and the approaches to counter it. These can be re-summarised into the following six points:

1.1 There are many forms of corruption – the scale and types of corruption, as well as the actors involved and the level of harm done, vary by context.

Corruption is employed by a multitude of actors in a multitude of circumstances and fora – from a politician receiving kickbacks from companies that have been awarded large contracts, to a teacher eliciting sex in return for good grades. Common forms of corruption include bribery, trading in influence, embezzlement, fraud, kickbacks, extortion, nepotism, clientelism, abuse of discretion, and graft. Determining which forms of corruption are most harmful is context specific. In some cases, many small acts of corruption are more harmful than one big act. In other cases, the opposite is true. The prevalence of each form of corruption varies both across and within countries – including by region, sector, and institution.

Corruption takes place within the political realm, the public sector, and the private sector – as well as intersections between the three. Corruption can also be trans-national and enabled by external factors such as tax havens and foreign policy objectives, including regional security, trade relations, and diplomacy. The ways in which external factors can enable corruption is clearly illustrated by Chayes' case study of the [structure of corruption in Azerbaijan](#).

In some contexts, corruption is an anomaly – the act of a few individuals. However, in many contexts, corruption is widespread. In its extreme, corruption is an integrated part of the socio-political and economic system. For a visual depiction of how complex and entrenched a system of corruption can be, see the Justice and Legitimacy Program's [systems map](#) of the dynamics driving police and judicial corruption in northern Uganda.

1.2 There are many anti-corruption tools, the details of which matter

There are numerous types of anti-corruption interventions. They can be global, regional, national, subnational, sectoral, or institution specific. They can be stand-alone or integrated into a larger programme. They can be explicit or

implicit; internal or external; direct or indirect. Many are not specifically designed with the prevention of corruption in mind, but rather for broader goals such as good governance or civil society strengthening.

Examples of common interventions that are funded in anti-corruption programmes include:

- Building the capacity of the police and judiciary
- Developing an e-procurement system
- Training investigative journalists
- Establishing, training, and funding an anti-corruption agency (ACA)
- Running anti-corruption advocacy campaigns
- Creating a community monitoring mechanism for a large health programme

A specific anti-corruption tool cannot be definitively labelled as being effective or ineffective – the details matter, as does the context of its implementation.

Several studies suggest that many conventional anti-corruption tools are ineffective.¹ However, a specific tool cannot be definitively labelled as being effective or ineffective – the details matter. Anti-corruption agencies, for example, vary in mandate, manpower, and degree of independence. Some have investigative and prosecutorial powers, while others do not. Some are spearheaded by strong individuals willing and able to indict politicians and their families, while others are not. These variations are key determinates of the likely impact that an ACA can have.² The context in which the tool is implemented – including the political settlement – also matters.

1.3 A unique combination of approaches, tools, and actors is needed to address the (often) multiple reinforcing drivers of corruption in a given context

Corruption is a complex phenomenon, often caused and sustained by multiple interlinked factors. These factors can be categorised into four drivers:

1. Rocha Menocal et al. 2015; Johnsn, Taxell and Zaum et al. 2012; Hanna et al. 2011; McGee and Gaventa 2010.

2. Johnsn et al. 2011; Schütte 2017.

principal–agent/institutional problems; collective action problems; justifying norms, values, and pressures; and the short-term functionality of corruption. These drivers are broad categories, and each requires further unpacking. A principal–agent problem could, for example, stem from a lack of transparency. This lack of transparency could, in turn, arise from the absence of an operational freedom of information law.

Anti-corruption interventions need to address the underlying drivers of corruption, not just the symptoms. In the past, the focus has largely been on ‘technical’ top-down reforms designed to address the principal–agent drivers of corruption.³ However, this approach is often insufficient. It can lead to a situation where, for example, income and asset declaration rules are put in place, but few people comply with them as they do not feel compelled to do so.

A problem-focused approach to understanding and addressing corruption requires exploring all the relevant stakeholders and drivers, as well as the wider socio-political and economic context. In any given context, the strength and combination of drivers, as well as the actors involved, will vary. Even when two countries, regions, sectors, or institutions appear to have the same corruption problem, the exact cause of it, as well as what solutions are feasible, will be circumstance specific. Effective tools that can be employed for such an exercise include Wedel’s Mapping Method and Marquette and Peiffer’s Corruption Functionality Framework.

Small, shallow, and isolated anti-corruption interventions are less likely to contribute to a sustained impact.

When multiple drivers of a corrupt act are present, a multi-pronged approach is likely needed. Small, shallow, and isolated anti-corruption interventions are less likely to contribute to a sustained impact.

1.4 When corruption is systemic, anti-corruption efforts need to take a systems approach

In some contexts, corruption is not only widespread, but also deeply entrenched. This phenomenon is often referred to as ‘systemic corruption’ or

3. Scharbatke-Church and Chigas 2016.

‘systems of corruption’. The [U4 Glossary](#) defines systemic corruption as: ‘A situation when corruption is an integral part of a state’s economic, social and political system, and where most people have no alternatives to dealing with corrupt officials.’ Jackson, Tobin and Eggert⁴ similarly state that: ‘Systemic corruption exists when a corrupt act recurs consistently and is connected to other corrupt acts through an underlying system that enables and encourages the corruption.’

While our understanding of how to address systemic corruption is still limited, it is clear that a different approach is likely needed. When corruption is systemic, anti-corruption interventions need to target the underlying drivers and the system itself, rather than symptoms or corrupt individuals. Convicting a powerful individual of corruption can be effective, if that sends a clear (and accurate) signal that the ‘rules of the game’ are changing. However, often the individuals that replace those that have been fired, transferred, convicted, etc due to corruption are (or become) corrupt themselves. Moreover, not everyone that engages in corruption is in favour of corruption. Many wish that the system would change but engage in corruption because of pressure or necessity. Construction contractors may, for example, risk bankruptcy and physical attack by their peers if they refuse to participate in well-established bid-rigging protocol.

When corruption is systemic, it can act like an equilibrium that is difficult to pierce.

When corruption is systemic, it can act like an equilibrium that is difficult to pierce. Anti-corruption efforts are most likely to be successful if they are (collectively) sufficient to bring about a new equilibrium. If they do not, the marginal effects of each intervention may diminish over time as vested interests find ways of undermining them and returning to the original equilibrium.⁵ There can be value in interventions that ‘lay the foundation’ for future reform efforts in contexts where the political settlement is not ready for a major anti-corruption drive. However, these interventions should be based on a strong theory of change and be in line with latest (and local) thinking on what interventions make sense given the context.⁶ If not, they may do more harm than good.

4. 2019.

5. Taylor 2018; Rothstein 2011; Fisman and Golden 2017a.

6. Mungiu-Pippidi 2016.

Interventions should be based on a strong theory of change and be in line with what make sense given the context – if not, they may do more harm than good.

For an overview of latest corruption theories and thinking, see Jackson.⁷

1.5 Anti-corruption interventions need to be flexible, politically responsive, and designed with potential backlashes in mind

Corruption is not a disease or deviation, but the historical baseline. Curbing corruption is therefore a complex, never-ending process. Progress is neither linear nor one-way. It is messy and uncertain. Even incremental improvements and smaller reforms can be difficult to bring about and sustain – especially within the timeframes of most development assistance projects.

Corruption and anti-corruption interventions are both political. Powerful and influential individuals and groups often have a vested interest in maintaining the status quo. The potential for anti-corruption interventions to encounter pushback and backlashes, or to bring about negative unintended consequences, is therefore ever present. Indeed, Fisman and Golden⁸ go as far as to state that ‘every anticorruption or antifraud program elicits a strategic response by those who orchestrated and benefited from wrongdoing in the first place.’

The following tables shows examples of potential negative unintended consequences and backlashes.

7. 2020.

8. 2017b.

Examples of potential negative unintended consequences and backlashes

Negative unintended consequences and backlashes	Examples
Corruption re-emerges in the same space but in a new form, or finds a way of playing with the new rules of the game.	<p>In order to remain anonymous under new requirements, beneficial owners in the extractive industries sector put in place nominee shareholders to represent them, or reduce their ownership level to below disclosure thresholds (Lemâitre 2019b; Lemâitre 2019a).</p> <p>In response to procurement reforms that make bidding more competitive, government buyers rely more heavily on non-competitive procedures types (Dávid-Barrett and Fazekas 2020).</p>
The form of corruption being targeted by the intervention not only remains intact but also increases.	Raising civil servant salaries results in an increase in petty corruption (Foltz & Opoku-Agyemang 2015).
The corrupt practice or actors move to a different sector, country, city, or institution (or arm of the institution).	A clampdown on corruption in one municipality leads to an increase in public-sector transfer requests to other municipalities.
The anti-corruption tool is misused or co-opted.	A new anti-corruption law or agency is used to prosecute and punish political rivals and critics (Johnston and Johnsen 2014).
The anti-corruption tools and/or their proponents are directly undermined or stopped.	A successful anti-corruption project is halted in the run up to an election due to pressure from those harmed by the intervention (Barnwal 2017; Fisman and Golden 2017a).
The broader development goal is undermined.	Removing the opportunity for frontline healthcare personnel to receive informal payments – without simultaneously addressing the challenge of very low wages – leads to a reduction in staff moral and an increase in healthcare personnel taking side jobs (Marquette and Peiffer n.d.; Peiffer et al. 2020; Mæstad and Mwisongo 2007).
The enabling environment for future anti-corruption reform is undermined.	<p>Anti-corruption projects and pledges with no, or limited, results leave citizens disillusioned (Johnston and Johnsen 2014).</p> <p>Anti-corruption messaging causes citizens to believe that corruption is so widespread and difficult to solve that they become apathetic towards it and lose faith in democracy (Cheeseman and Peiffer 2020).</p>

The likelihood of unintended consequences and backlashes can be mitigated by designing anti-corruption interventions based on a strong theory of change and

a clear understanding of the political economy and political settlement. Key questions that need to be answered include: ‘Which powerful stakeholders are for and against reform and why?’ and ‘How might the system “push back” against efforts at reform?’.⁹

Many academics and practitioners believe that anti-corruption interventions should target corruption problems that are feasible to address, given the political settlement.¹⁰ Approaches such as taking advantage of windows of opportunity, building political will and trust, changing expectations, collaborating and coordinating with others, proactively working to maintain momentum, and continuously monitoring and adjusting interventions also show promise.¹¹

Finally, the level and genuineness of stakeholders’ commitment to anti-corruption efforts should not be taken at face value. Politicians may, for example, need to appear to be making progress on good governance in order to appease voters or receive international aid.

1.6 There are limits to what anti-corruption interventions alone can achieve, as well as the role that donor agencies can play, so successful anti-corruption efforts may require a broader approach

There are limits to what even the perfectly designed and implemented anti-corruption interventions can achieve. Training investigative journalists to uncover corruption, for example, can lead to an increase in acts of corruption being brought to light. However, whether this will lead to an increase in arrests, prosecutions, and convictions depends on a range of factors beyond the control of that intervention.

Similarly, there are limits to the role that external actors can play. Countries that have significantly reduced their corruption levels have primarily done so due to internal forces. Klitgaard¹² finds that ‘providers of development assistance can contribute resources, knowledge, convening power and leverage that may help recipient countries reform.’ However, political will and bottom-up

9. Scharbatke-Church and Chigas 2016.

10. Khan, Roy and Andreoni 2016; Uberti 2020; Taylor 2018; Heywood 2018; Levy 2014.

11. Wathne 2021.

12. 2015.

pressure are essential. Mason¹³ believes that the role of external pressure may also be enhanced if countries employ a whole-of-government approach, rather than limit anti-corruption measures to development assistance.

Part 2: Implications for anti-corruption evaluations

The nature of corruption and the current approach to anti-corruption initiatives have a number of implications for evaluations – from the way data are gathered to the way sustainability is assessed. The remainder of this Issue aims to explain these theme-specific challenges and offer potential ways to overcome them. While the general messages will be familiar to most readers, the specific issues relating to evaluating anti-corruption interventions may not.

In addition to these guidelines, evaluations should adhere to evolving standards of best practice. This entails, amongst other things, carrying out evaluations that are:

- Timely, useful, and used¹⁴
- Clear, focused, and tailored
- Genuinely¹⁵ participatory¹⁶
- Gender-, human rights-, and sustainability-responsive¹⁷
- In adherence with quality and ethical standards, as well as aid effectiveness principles¹⁸

Many evaluations are commissioned and organised in accordance with the six OECD/DAC evaluation criteria. The following table summarises the applicability of the paper's main messages to each of these. It can be used to guide readers to the sub-sections most relevant for their needs.

13. 2020.

14. Patton 2012; Johnsen et al. 2011.

15. The term 'genuine' is included to emphasise the need for real participation and consultation, as opposed to tokenism or window dressing. This distinction is starkly illustrated by Roger Hart's 1992 eight rung ladder of participation.

16. South to South Initiative.

17. UNEG 2016; Prague Declaration on Evaluation for Transformational Change; Raimondo and Bamberger 2015; Blue Marble Evaluation; Fletcher 2015; Espinosa 2013.

18. Sida 2020.

Employ a realistic barometer of intervention success – even if this means questioning the logframe

Applicability to each OECD/DAC evaluation criterion

Message	Relevance	Coherence	Effectiveness	Efficiency	Impact	Sustainability
Consider the need for a complexity-responsive evaluation methodology			X		X	X
Carefully gather and make use of data			X	X	X	X
Employ a realistic barometer of intervention success – even if this means questioning the logframe			X		X	
When attribution cannot be determined, aim to demonstrate contribution			X		X	
Assess the extent to which there is a problem-strategy match	X		X		X	X
Determine whether the intervention is in line with latest and local thinking on corruption and anti-corruption efforts	X	X	X		X	X
Explore the potential unintended consequences	X		X		X	X

Message	Relevance	Coherence	Effectiveness	Efficiency	Impact	Sustainability
and the likelihood of a backlash						
Draw on all of these insights, and more, to assess sustainability						X
Determine the impact and 'what works', using appropriate evaluation methods and tools					X	
Assess the effectiveness and impact at the appropriate level			X		X	

When attribution cannot be determined, aim to demonstrate contribution

2.1 Consider the need for a complexity-responsive evaluation methodology

As explored in Part One, corruption and measures to curb it are often complex. Bamberger, Vaessen, and Raimondo¹⁹ identify five, often interlinked, dimensions of potential complexity in development evaluations: 1) the intervention; 2) the institutions and stakeholders involved in the intervention; 3) the effects of the intervention – including the causal change processes; 4) the context in which the intervention is implemented; and 5) the evaluation itself.

The extent to which complexity needs to be considered when evaluating anti-corruption interventions will vary. An intervention may be more complex in some dimensions than in others. A checklist for assessing the level and areas of complexity has been developed by Bamberger, Vaessen, and Raimondo.²⁰ The

19. 2015.

20. 2015.

following table adapts this checklist in order to highlight areas of potential high complexity.

Assessing the level and areas of complexity

<p style="text-align: center;">Dimension 1: The nature of the intervention</p> <ul style="list-style-type: none"> • Are there multiple, broad, and/or unclearly defined objectives? • Does it affect a large population? • Is the programme design emergent? • Are implementation procedures unclear or changing? • Are there many services or components? • Is there high technical complexity? • Is there high social complexity? • Does it lack a clear start or end date? • Is the programme design relatively new or untested? 	<p style="text-align: center;">Dimension 2: Institutions and stakeholders</p> <ul style="list-style-type: none"> • Is it funded with general budget support with no clear definition of the services to be funded? • Are there many funding and implementing agencies involved? • Are there many and/or diverse stakeholders?
<p style="text-align: center;">Dimension 3: Causality and change</p> <ul style="list-style-type: none"> • Are there multiple causal pathways? • Is there a low degree of certainty regarding outcomes? • Is there a low level of agreement and clarity regarding how to address problems? 	<p style="text-align: center;">Dimension 4: Embeddedness and nature of the system</p> <ul style="list-style-type: none"> • Are the contextual factors and their potential influence on implementation and change largely unknown? • Are there multiple mechanism to promote complex behavioural change?

(Source: Adapted from the table 'Checklist for assessing levels of complexity' in Bamberger, Vaessen, and Raimondo (2015))

When a complexity-responsive evaluation is deemed necessary, evaluators may need to complement established evaluation approaches with emerging methods – such as methods from complexity science or innovative data-collection techniques.²¹ Several potential methods are already becoming more mainstream. These methods include outcome harvesting, qualitative comparative analysis, process tracing, contribution analysis, social network analysis, systems mapping, and appreciative enquiry.

For an overview of complex collection and analysis methodologies that may be relevant when evaluating anti-corruption interventions, see INTRAC.²² Similarly, for a framework for determining which impact evaluation methods are most appropriate when evaluating the complex, see section 2.9.

21. Raimondo, Vaessen and Bamberger 2015.

22. 2017.

It may also prove useful to consider evaluation criteria beyond those defined by OECD/DAC, as shown in Box 1.

Box 1: Going beyond the OECD/DAC criteria to assess interventions addressing complex problems

Most evaluations are based on a sub-set of the six OECD/DAC evaluation criteria. However, the recent evaluation of Norway's anti-corruption efforts saw the need to explore two more: flexibility and learning. The criteria draw on the Problem-Driven Iterative Adaptation (PDIA) approach, which recognizes that solving complex problems – such as corruption – requires a flexible approach that takes into account continual learning.

(Source: Vaillant et al. 2020)

2.2 Carefully gather and make use of data

Evaluators of anti-corruption interventions often face four data challenges: getting accurate data; getting intervention-specific data; getting sufficient data; and securing data in an ethically responsible and safe manner.

Corruption is clandestine and illegal. It is therefore difficult to measure fully and directly.²³ As a result, evaluators must often rely on imperfect estimates such as proxy indicators, perception surveys, and expert opinion. At best, these sources of data provide a realistic estimate of corruption and its consequences. However, their findings can also be misleading. For example, a crackdown on corruption can result in the public becoming more aware of the extent of corruption. This can lead to a worsening in perception-based corruption survey results and, potentially, public apathy towards corruption.

The available data may also not be intervention specific. Many interventions' logframes and monitoring systems rely in part on secondary data. These data are often too aggregated to be attributable to the intervention itself or to capture small changes. Global indexes that can be disaggregated by type of corruption, sector, or institution are generally preferable to global indexes with no basis for disaggregation. However, even these are often beyond the intervention's sphere

23. Dávid-Barrett et al. 2020; Mugellini, Villeneuve and Heide 2021.

of influence, as further explained in section 2.4 on attribution. Locally generated data, such as administrative statistics and target surveys, are more likely to reflect a specific intervention's effects.²⁴

While programme-specific indicators are the best gauge of an intervention's performance, not all anti-corruption interventions have sufficient baseline and monitoring data.²⁵ This is particularly common when the anti-corruption intervention is implicit or mainstreamed. Because anti-corruption efforts are political in nature, the actors behind the intervention may choose a clandestine approach and decide not to rigorously document results. As Robillard and Robillard²⁶ point out, making an anti-corruption intervention explicit 'makes it easy for people in power to applaud these initiatives in public – and to avoid them, or even undermine them, in private. By the time the project reports are written, the systems that facilitate corruption will have shifted, adapted, and survived.'

The political nature of corruption also means that stakeholders may be unwilling to participate in interviews or other data-gathering methods. Even if they do, they may hold back useful information for fear of sanctions or retaliation.²⁷ Such concerns are justified, as whistleblowers, journalists, and reformists alike have lost their jobs, been arrested, had their reputations destroyed, been threatened, and – in some cases – even killed.²⁸ Gaining access to, and eliciting honest answers from, the individuals facilitating and engaging in corruption can also be both difficult and ethically complex.

A final challenge is that there is no universally agreed definition of corruption.²⁹ While studies have found that people's view of what corruption comprises is relatively similar across societies,³⁰ there are differences in opinion. Small informal payments may not always be viewed as a bribe. Similarly, corruption that benefits the community or family can be considered morally legitimate.³¹

In response to these challenges, the commissioners of anti-corruption evaluations should:

24. Hart 2019.

25. Winbourne and Spector 2014; ICAI 2014; Spector, Winbourne and Dininio 2015.

26. 2018.

27. Council of Europe 2013, p. 14.

28. U4 2017.

29. Mugellini, Villeneuve and Heide 2021.

30. Rothstein and Varraich 2017, p. 47; Kurer 2015, p. 38.

31. Smith 2015.

- Determine the evaluability of the intervention at the outset³²
- Select a diverse evaluation team that brings together evaluation, political economy, thematic, and local expertise, as well as access to stakeholders
- Consider any biases, pressures, or repercussions that they or the evaluation team may face

Similarly, evaluators should:

- Prioritise the safety and comfort of evaluation participants
- Carefully choose their words to ensure clarity and avoid unnecessary controversy
- Use data-gathering methods that allow for more open and honest input (eg anonymous surveys, informal talks, vignettes, etc)
- Tailor the choice of data-collection approach to the type of corruption being discussed and the stakeholder being questioned³³
- Gather data from a range of stakeholders with different perspectives
- Base their analysis on primary and secondary data that are attributable to the intervention
- Use a mixed-method approach and triangulate findings
- Consider the quality of the various data sources³⁴

Evaluators should also take into account standard data-collection issues, such as how to overcome social desirability bias and discern perceptions from reality.³⁵ The team should also disaggregate data by gender, poverty, and other relevant socioeconomic and demographic characteristics as corruption, and measures to counter it, can impact – directly and indirectly – groups in different ways.³⁶

For general advice on how to address data constraints, including missing or difficult to collect data, see chapter five of the third edition (2020) of Bamberger and Mabry's book *RealWorld evaluation: Working under budget, time, data, and political constraints*. For an assessment of the strengths and weakness of potential corruption measurement tools see Hart,³⁷ Trapnell, Jenkins and

32. Page 7 of DfID's *Evaluation Strategy 2014–2019* provides a simple checklist to help determine whether to carry out an evaluation.

33. Scharbatke-Church with Barnard-Webster 2017.

34. Mugellini, Villeneuve and Heide 2021.

35. Scharbatke-Church, Barnard-Webster and Woodrow 2017.

36. Kirya n.d.

37. 2019. See also the accompanying annexes from *U4 Guide 2019:1: Measurement and assessment tools table and Reference table on matching measurements and assessment tools to corruption diagnostic questions*.

Chêne,³⁸ and Trapnell.³⁹ For further details on data generation tools and approaches that are well suited to sensitive topics such as corruption, see Scharbatke-Church,⁴⁰ as well as the section on methods in Jackson and Köbis,⁴¹ pages 26–29 of Scharbatke-Church with Barnard-Webster,⁴² and sections 3.1.2 and 3.1.3 of the Council of Europe.⁴³ For insights on how to measure changes in social norms, see Scharbatke-Church and Kothari.⁴⁴ See also Box 2 for lessons learned from studying corruption in the criminal justice system.

Box 2: Data gathering: Lessons learned from studying corruption in the criminal justice system

- Duty bearers and rights holders have a different experience and understanding of the same system of corruption.
- While informed consent is needed, a formal consent letter may not be appropriate and could intimidate participants wishing to ensure their anonymity.
- Local intermediaries that are known and trusted can help the evaluation team reach out to stakeholders and secure authentic answers.
- Interviews should be conducted in locations that allow participants to speak honestly and without fear of reprisal.
- Some forms of corruption – such as sexual favours – are more taboo to talk about than others. Which topics interviewees are willing to discuss also vary according to who is asking the questions.
- Participatory data-collection processes such as interactive discussions can generate important insights. However, these methods may be more difficult to use when there are power dynamics in the room, or sensitivities regarding blame.
- Vignettes work well in focus groups of average citizens, but not with those involving police. With that particular stakeholder, vignettes can still be used to facilitate more honest and free discussions, but only during one-to-one interviews.

(Source: Scharbatke-Church with Barnard-Webster 2017)

38. 2017.

39. 2015.

40. 2017.

41. 2018.

42. 2017.

43. 2013.

44. 2021.

2.3 Employ a realistic barometer of intervention success – even if this means questioning the logframe

As discussed in section 2.2, anti-corruption interventions do not always have clearly stated goals, objectives, and indicators with accompanying baseline and monitoring data. Even when they do, these indicators are not always a fair barometer of success. Some anti-corruption interventions have logframes with indicators outside their sphere of control and influence.⁴⁵ A country's score or rank on the Transparency International Corruption Perceptions Index is, for example, not an accurate gauge of the impact of establishing and building the capacity of an anti-corruption agency, regardless of how effective the agency may be. Less ambitious indicators – such as the number of prosecutions – are also often dependent on a range of factors beyond an intervention's sphere of control or influence.

When assessing effectiveness and impact, evaluators must keep in mind that even small changes in corruption levels are difficult to bring about and sustain. The impact of most interventions will be modest in scale⁴⁶ and a project cycle is often too short a time period to bring about significant and lasting change.

In contexts where anti-corruption efforts are being actively undermined, preventing an increase in corruption or holding on to past gains may in itself be considered a success.

An assessment of effectiveness and impact also needs to consider the context. It is, for example, more difficult to reduce corruption in fragile states and to address corruption that is ⁴⁷ In contexts where anti-corruption efforts are being actively undermined, preventing an increase in corruption or holding on to past gains may in itself be considered a success.

45. Winbourne and Spector 2014.

46. World Bank 2020.

47. Rothstein and Tannenbergh 2015, p. 68.

2.4 When attribution cannot be determined, aim to demonstrate contribution

Demonstrating the causal link between a specific anti-corruption intervention and concrete outcomes and impacts can prove difficult – particularly at the higher level.⁴⁸ First, multiple donors and interventions may be supporting the same target unit (eg the judiciary). Second, change is often brought about by a multitude of past and present forces. Third, there are often data limitations. In many cases, it is therefore appropriate to aim to demonstrate contribution rather than attribution.

Imagine, for example, that leaders of a corruption scheme are found out, prosecuted, and convicted. This ‘impact’ is potentially the result of training investigative journalists (who first brought the case to light); getting a freedom of information law passed (so that the journalists were able to access evidence); providing the police, prosecutors, and judges with capacity building (so that a solid case could be built and fairly tried); a newly elected prime minister wishing to demonstrate that she is tough on crime (ensuring the necessary political will); and corruption awareness campaigns (creating mounting pressure to act).

There are four approaches to establishing a causal inference: counterfactual frameworks; regularity frameworks; configurational frameworks; and generative frameworks.⁴⁹ Counterfactual methods, such as randomised control trials, are often touted as the preferred approach. However, the value of other approaches is increasingly recognised. Specific methods include RAPID Outcome Assessment (ROA), episode studies, and contribution analysis. The first two methods begin with the change itself and then map the process backwards to determine causal factors. This can be an effective way to ensure that the importance of the intervention is not overestimated. Triangulation can also help overcome attribution and contribution challenges.⁵⁰

48. Hart 2019; Dávid-Barrett et al. 2020.

49. Befani 2012; Stern et al. 2012; Befani and Mayne 2014.

50. OECD 2021.

2.5 Assess the extent to which there is a problem–strategy match

The OECD/DAC criteria relevance does more than determine whether an intervention is in line with local and global needs, priorities, and policies. It aids in recognising the likelihood of effectiveness, impact, and sustainability. The OECD guidance states that the term ‘responds to’ in the definition of relevance includes whether ‘the objectives and design of the intervention are sensitive to the economic, environmental, equity, social, political economy and capacity conditions in which it takes place.’

An intervention can address the ‘right’ problem, but in the ‘wrong’ way.

An intervention can address the ‘right’ problem, but in the ‘wrong’ way. By analysing whether there is a problem–strategy match, evaluators can get an indication of how likely an anti-corruption intervention is to achieve results and be sustainable, as well as avoid negative unintended consequences (a topic explored in section 2.7). Unfortunately, anti-corruption interventions often lack an explicit and clear theory of change.⁵¹ However, when they do, evaluators can compare this theory of change to the situational analysis, including any recent updates. They can also work with stakeholders to retroactively create a theory of change.

A key question to explore is whether the intervention was appropriate, given the context. Even when the same corruption problem appears in two different contexts, the same solution may not work in both, as the underlying drivers and wider conditions will vary. Hanna et al.,⁵² for example, find that community-level monitoring can be an effective anti-corruption tool, but only when the community is able to punish corruption. Such insights have implications both for maximising results and doing no harm. As Johnston states, ‘What might seem to be a good reform idea in country A may well be impossible in B, irrelevant in C, and downright harmful in D’.⁵³

Another important question is whether the intervention, either on its own or in combination with other efforts, is sufficient to address the problem – and,

51. Scharbatke-Church and Chigas 2016; Johnsen 2012; Taylor 2018.

52. 2011.

53. 2014 p. 3.

ideally, pierce the corruption equilibrium. Why? Because small, shallow, and isolated anti-corruption interventions are less likely to contribute to significant changes and more likely to have their effect diminished over time.

2.6 Determine whether the intervention is in line with latest and local thinking on corruption and anti-corruption efforts

Relevance, effectiveness, impact, and sustainability are in part determined by whether the intervention was (and remains) designed and implemented in line with latest and local thinking on corruption and anti-corruption measures. Three decades of anti-corruption efforts have not delivered the anticipated level of results.⁵⁴ There are many justifiable reasons for this – including the difficulty of combating corruption. However, the experience to date can in some degree be attributable to weaknesses in conventional anti-corruption approaches.

Part One of this Issue summarises many of the key lessons learned. The reader is particularly encouraged to familiarise themselves with sub-sections 1.3, 1.4, and 1.5.

Other anti-corruption resources that can be drawn on to determine whether interventions are likely to succeed include: Jackson 2020; World Bank 2020; Mungiu-Pippidi 2017; Rocha Menocal et al. 2015; Johnsen, Taxell, and Zaum 2012; Hanna et al. 2011; and Anti-Corruption Evidence (ACE) Research Consortium 2019.

Some of these publications focus exclusively on the overarching principles – such as the need to address all drivers of corruption, and not just principal–agent problems. Others assess the evidence base for specific anti-corruption tools. While evidence on ‘what works’ is limited, it is growing. For example, Johnsen, Taxell, and Zaum⁵⁵ and Rocha Menocal et al.⁵⁶ find that public financial management can be an effective anti-corruption tool. Similarly, although universally applicable formulas can never be created, evidence is emerging on potentially complementary approaches. For example, Mungiu-

54. Scharbatke-Church with Barnard-Webster 2017; ICAI 2014; Mungiu-Pippidi 2015; Heeks 2011; Persson, Rothstein, and Teorell 2019; Khan, Roy and Andreoni 2019.

55. 2012.

56. 2015.

Pippidi and Dadašov⁵⁷ find that anti-corruption legislation is more effective when there is freedom of the press and independence of the media.

While best practice is one important source of knowledge, it is not the only source. Local knowledge on what has worked, what has not worked, what is feasible, and why, is equally insightful. See section 4 of the Problem-Driven Iterative Adaptation (PDIA) toolkit by the Building State Capability Program for details.

2.7 Explore potential unintended consequences and the likelihood of a backlash

As discussed in section 1.5, the likelihood of anti-corruption interventions bringing about unintended consequences and backlashes being experienced is high. Unintended consequences can be positive or negative. Similarly, a backlash can be immediate or years in the making. The Indonesian Corruption Eradication Commission (KPK), for example, successfully staved off efforts to undermine its credibility, independence, and effectiveness for over 20 years. However, in 2019 the House of Representatives quickly pushed through legislation that significantly reduced the agency's autonomy.⁵⁸

Evaluators should proactively seek to identify unintended consequences and backlashes that have already transpired or are likely to happen in the future. Potential ways of doing so include:

- Assessing the extent to which there is a problem–strategy match (see section 2.5)
- Speaking to third party experts with a deep understanding of the local context and ‘how things work’ – eg grassroots leaders, anthropologists, etc.
- Selecting evaluation tools designed to uncover unexpected results – eg Most Significant Change (MSC) stories and tracer studies
- Exploring the current political settlement, including whether there is sufficient political will to sustain reform efforts (see section 2.8)

For additional guidance on how to incorporate a focus on unintended consequences into an evaluation, see Bamberger, Tarsilla and Hess-Biber.⁵⁹

57. 2017.

58. Schütte 2019.

59. 2016.

2.8 Draw on all of these insights, and more, to assess sustainability

According to the OECD/DAC revised evaluation criteria, an assessment of sustainability ‘[i]ncludes an examination of the financial, economic, social, environmental, and institutional capacities of the systems needed to sustain net benefits over time. [... It involves] analyses of resilience, risks, and potential trade-offs.’

Like all interventions, an analysis of the sustainability of an anti-corruption intervention entails asking questions such as ‘Does the intervention have sufficient funds to continue?’ and ‘Are the stakeholders involved in the intervention interested in continuing to be active?’ Many community monitoring groups, for example, become inactive once funding for the programme that supported them comes to an end.

However, there is also another level of questions that need to be answered when assessing how likely results are to be sustained in the long term. For example:

- Did the intervention address the underlying drivers of corruption?
- Has the gap between the formal and informal rules of the game been closed?⁶⁰
- Is the intervention backed by powerful and influential stakeholders?⁶¹
- Does the political settlement support these reforms?⁶²
- Has the intervention (in combination with other reform efforts) succeeded in shifting the landscape to a new corruption equilibrium?⁶³

For further details, see section 2.5 on the need for a problem–strategy match and section 2.7 on backlashes. For additional guidance on how to incorporate sustainability questions into evaluations see Chelimsky.⁶⁴

60. Jackson and Köbis 2018.

61. World Bank 2020, p. 343; Mungiu-Pippidi 2016.

62. Khan, Roy and Andreoni 2019.

63. Taylor 2018; Jackson 2020.

64. 2019.

2.9 Determine impact and ‘what works’, using appropriate anti-corruption methods

Additional studies that rigorously assess impact can help fill the anti-corruption evidence gap, as there is insufficient academic research on ‘what works and why’.⁶⁵ If an evaluation is going to assess impact, it should do so thoroughly. A decade ago, few evaluations of anti-corruption interventions focused on impact. Those that did generally made use of insufficiently rigorous methodologies.⁶⁶ As background research for this Issue, we reviewed 37 anti-corruption and good governance evaluations conducted between 2014 and 2020. While 16 of the 37 evaluations claim to assess impact, very few do. Most of the methodologies utilised are also not suitable for measuring impact. These findings are in line with the results of the industry-wide consultation survey carried out as part of the OECD/DAC criteria revision process: only 34% of respondents were satisfied with the way the ‘impact’ criterion is implemented – compared to a satisfaction rate of over 75% for the ‘relevance’ and ‘effectiveness’ criteria.⁶⁷

When evaluating the impact of an intervention, it is important to be issues-driven, rather than methods-driven.⁶⁸ The evaluation approach, methodology, and techniques should be guided by the overall purpose and scope of the evaluation; the questions that need to be answered; the intervention’s attributes; the context; the availability of data; and the allotted budget, time, and human resources. With regards to the field-specific considerations, Wathne⁶⁹ states: ‘Rather than a binary analysis of their impact – as effective or not – we need to understand whether, why, to what extent, under what circumstances, in which contexts, in which combinations, and for whom anti-corruption efforts have a direct or indirect impact on corruption levels and, ultimately, on development outcomes.’

This statement acknowledges that:

- What works in one situation, may not work in another.
- There are multiple pathways to bringing about a reduction in corruption.
- A specific anti-corruption tool is unlikely to be sufficient on its own. A multi-pronged approach is generally needed. Using logic terminology, the implied

65. Hanna, et al. 2011; Johnsen, Taxell and Zaum 2012; Rocha Menocal et al. 2015.

66. Johnsen, Taxell and Zaum 2012; Johnsen and Søreide 2013; Hanna et al. 2011.

67. OECD 2018.

68. Vaessen, Raimondo, Bamberger 2015.

69. 2021.

impact question becomes: ‘Which anti-corruption tools are “necessary but not sufficient” vs. “neither sufficient nor necessary”?’

- Quantifying the impact of anti-corruption interventions is challenging, given the difficulty of measuring corruption and changes in corruption⁷⁰ as well as determining an intervention’s level of contribution.
- Partial successes and limited results are often all that are feasible within a given budget, time frame, or context.
- Determining causality is difficult.⁷¹
- Progress is not linear. Long periods with no results can be followed by rapid change, and setbacks are likely. Gauging the likely impact trajectory over time may therefore be as important as documenting the level of impact at a given point in time.⁷²
- Interventions can be multi-pronged and bring about multiple results; bring about different results for different stakeholders;⁷³ and lead to unanticipated results and consequences.
- The details of a specific anti-corruption tool matter. As explored in section 1.2, anti-corruption agencies vary considerably in terms of their mandate, manpower, degree of independence, and the context in which they operate. This variation goes a long way to explaining why some ACAs have been more effective than others.

Evaluators need to design an impact evaluation that considers the above realities of corruption and anti-corruption measures. The ability of each evaluation method to gauge impact in the context of complexity can be assessed along six dimensions:⁷⁴

- **Attribution:** Documenting changes and attributing them to the intervention
- **Explanation:** Showing how the intervention worked and how it affected change
- **Multiple causal pathways:** Exploring multiple changes and/or the confluence of factors impacting change
- **Nature of causal change:** Assessing change that is uncertain, non-linear, and emergent
- **Emergence:** Accommodating both the intervention adapting over time and

70. Johnsen, Taxell and Zaum, 2012; World Bank 2020.

71. Johnsen, Taxell and Zaum 2012; Rocha Menocal et al. 2015.

72. Woolcock 2013.

73. Hart 2019.

74. Vaassen, Raimondo, and Barberger, 2015.

the wider context changing over time

- **Scope of effects:** Capturing all effects, including unintended outcomes

A number of evaluation handbooks have clear statements regarding which methods are best when determining impact and ‘what works’. Many of these handbooks favour quantitative methods. The World Bank/IDB 2016 handbook, *Impact Evaluation in Practice*, for example, includes five chapters on quantitative design approaches. Reference to qualitative approaches is limited to a sub-section on complementary approaches, where they are referred to as valuable tools for supplementing findings. Similarly, 3ie’s ‘Impact evaluation glossary’ defines an impact evaluation as having ‘either an experimental or quasi-experimental design.’ The glossary goes on to define a small ‘n’ impact evaluation as ‘the set of best available methods when “n” is too small to apply statistical approaches to constructing a counterfactual.’ However, qualitative approaches are not just a means of supplementing and triangulating quantitative findings. They are important approaches in their own right. Qualitative approaches are, for example, generally better placed to explore unintended consequences.⁷⁵

For an overview of qualitative methods for assessing impact, see the DfID’s 2012 Working Paper. In it, the authors show how and when theory-based, case-based, and participatory approaches can be the preferred choice for an impact evaluation. For an overview of quantitative methods, see Johnsen and Søreide⁷⁶ and Garcia.⁷⁷

The international development network Bond has developed an open-access tool to help evaluators choose the right method. Examples of when various methods may be appropriate are also provided in Box 3. In most cases, a genuinely mixed methods approach is best⁷⁸ – a topic explored further in Box 4.

75. Roelen and Devereux 2014.

76. 2013.

77. 2011.

78. Bamberger 2015.

Box 3: A sample of the range of methods that can be employed to evaluate anticorruption interventions

Process tracing: Process tracing could be used to explore why new anti-corruption legislation was passed, and what role a particular anti-corruption intervention may have had in bringing about this result. The ANTICORRP project included a number of process tracing studies, including a study of recent anti-corruption reforms in Taiwan (Göbel 2015) and a study of the ‘tipping points’ that limited the possibility for corruption in Costa Rica (Wilson and Fernández 2015).

Qualitative comparative analysis (QCA):QCA could be used to explore which traits and contexts enable anti-corruption agencies to successfully combat corruption. QCA was used by Stevens (2016) to explore potential causal conditions of national levels of corruption.

Randomised control trial (RCT): An RCT could help assess the impact of a specific tool or training that is rolled out in phases over a large number of units – for example video cameras in highway toll stations (Johnsøn and Søreide 2013). A randomised field experiment was used by Olken (2005) to compare the impact of two anti-corruption tools in road construction projects. More recently, a randomised control trial was included within a social accountability programme in Uganda (Fiala and Premand 2018).

Two common challenges when evaluating anti-corruption interventions are the sample size and approach. Anti-corruption interventions can have a target unit ranging from one (eg increasing the integrity of an agency) to millions (eg running a national campaign to raise awareness of corruption). However, the ‘n size’ is often small. Interventions also often consist of multiple activities and contribute to multiple results. Assessing impact is still possible. Johnsøn and Søreide⁷⁹ propose three approaches to evaluating anti-corruption interventions with a small target unit and/or multiple components: employ theory-based evaluation tools; use a formative (real-time) evaluation; or break an intervention into its constituent parts and evaluate each work stream separately. In so doing, it is important to keep in mind that the overall impact of an intervention can be more than the sum of its parts. The process of reassembling must therefore capture interactions between the work streams.⁸⁰

79. 2013.

80. Bamberger, Raimondo and Vaessen 2015.

The rise of big data, machine learning, and artificial intelligence may also open up new opportunities for evaluating impact. Governments are increasingly adopting e-government and open data projects. These data could potentially be analysed over time and across jurisdictions. See Berliner and Dupuy;⁸¹ White;⁸² van de Braak, Choenni and Bamberger;⁸³ and Leouze, Areias and Jackson⁸⁴ for details.

Box 4: A tailored, mixed-methods approach to evaluating anti-corruption agencies

Dávid-Barrett et al. (2020) recently conducted two theory-based evaluations to assess donor support to anti-corruption agencies. The evaluators' headline message resonates with this Issue: 'Evaluating anti-corruption programmes is difficult because of the complexity and hidden nature of corruption, its political sensitivity, and the ability of corrupt networks to adapt.'

The evaluators chose to focus on intermediate outcomes – rather than long-term goals. These were mid-programme evaluations, not impact evaluations. They were also conscious of the time and resources that it takes to achieve higher-level results, such as reducing corruption and catching and prosecuting 'big fish'.

Three tailored tools to evaluate ACAs were employed: an organisational capacity assessment, a network analysis, and policy tracking. The capacity assessment tool was in part based on existing tools. However, the choice of criteria also drew from a theory of change exercise with stakeholders that allowed the evaluators to identify the kinds of capacity that were particularly important in these contexts.

The network analysis tool analysed the agencies' formal and informal networks in order to determine whether they had the coordination, collective action, and support necessary to carry out investigations and develop cases. The policy tracker was used to gauge the extent to which agencies were able to advance policy and advocacy goals, such as closing regulatory loopholes that create opportunity for corruption. The tool was selected in recognition that the long-term success of these agencies depends upon the wider context.

81. 2018.

82. 2019.

83. 2015.

84. 2015.

2.10 Assess effectiveness and impact at the appropriate level

In addition to deciding the most appropriate design and method, the level of results to be evaluated needs to be clear. Is it, for example, sufficient to determine changes in capacity or attitudes? Or should the evaluation aim to determine changes in practice, such as asset declarations being filled out and reviewed, or an increase in corruption convictions? Alternatively, should the evaluation determine whether and by how much the level of corruption has been reduced and whether this has caused a significant change in the lives of beneficiaries? An anti-corruption intervention in the education sector may, for example, aim to contribute to increasing students' learning by ensuring that funds allocated for books, desks, and classrooms are properly spent and that teachers regularly show up and teach the full curriculum.

Alternatively (or additionally), the evaluation can focus on the extent to which underlying drivers of corruption have been addressed – whether, for example, incentives and social norms have changed. This distinction builds on material produced by the author, together with David Jackson, that suggests that anti-corruption interventions which target systems of corruption can theoretically have an impact at three levels: reducing or eliminating a specific corrupt practice; weakening the larger system of corruption; and contributing to nurturing deeper systems of accountability in a society.

The final two levels of impact are in line with current calls⁸⁵ for an evaluation approach that is appropriate for assessing transformational change and systems change. As Patton⁸⁶ says: 'Increasing the number of beds in shelters for the homeless constitutes increased impact. Changing the housing system so that no one is homeless is transformational.' Burns⁸⁷ similarly points out that 'If our aim is to create sustainable change, we have to show more than that an intervention created a change. We have to show that the system dynamic changed. [... It is therefore necessary to] focus our impact assessment on how the system dynamic has changed, not on what changes have taken place within the system dynamic.'

85. The IDEAS [Prague Declaration on Evaluation for Transformational Change](#) commits all members to 'evaluating for social, environmental and economic sustainability and transformation' in all evaluations – regardless of the programme or policy being studied.

86. 2020.

87. 2014.

The revised OECD/DAC impact criteria acknowledge this necessity. A suggested evaluation question in the [OECD guidance](#) is therefore: ‘Is the intervention transformative – does it create enduring changes in norms – including gender norms – and systems, whether intended or not?’ The guidance goes on to state the following with regards to transformation: ‘The definition defines transformational change as “holistic and enduring changes in systems or norms”. Transformational change can be thought of as addressing root causes, or systemic drivers of poverty, inequalities, exclusion and environmental damage, and is recognised in the 2030 Agenda as necessary to achieving the sustainable development goals.’

A focus on transformational changes entails assessing significance, depth of change, and scale of change. For further details see Feinstein,⁸⁸ the Independent Evaluation Group at the World Bank⁸⁹ and the [Blue Marble Evaluation](#).

Filling the anti-corruption evidence gap – the benefits of effective evaluations

The findings from the [2021 Transparency International Corruption Perception Index](#) are sobering: the global average score (43 out of 100 points) remains unchanged for the tenth consecutive year. While this aggregate figure masks local successes, the findings confirm that corruption remains widespread.

Evaluations can help policymakers and practitioners to better design and implement anti-corruption strategies and interventions, but only if the evaluations are both useful and used. This is by no means a revolutionary conclusion. Indeed, most monitoring and evaluation handbooks emphasise the importance of a system that facilitates both accountability and learning. Yet, in practice, the focus is often still on accountability. There is also still a tendency to highlight and share successes, while glossing over challenges and failures.

When evaluating, there is still a tendency to highlight and share successes, while glossing over challenges and failures –we need to raise the evaluation bar.

88. 2019.

89. 2016.

We need to raise the evaluation bar. This requires planning for evaluations from the outset so that there are sufficient data to carry out a quality evaluation; designing and undertaking evaluations using methods and tools that are tailored to studying the problem and the intervention (in this case, corruption and the measures to curb it); and, lastly, openly sharing and actively disseminating the findings from evaluations in order to facilitate learning and improve future policies and practice.

References

3ie. 2012. Impact evaluation glossary.

Anti-Corruption Evidence (ACE) Research Consortium. 2019. Anti-corruption in adverse contexts: Strategies to improve implementation. Briefing paper 006.

African Evaluation Association. n.d. South to South Initiative (S2SE).

Bamberger, M. 2015. *The importance of a mixed methods approach for evaluating complexity*. In Dealing with complexity in development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 144–164. SAGE.

Bamberger, M. and Mabry, L. 2020. RealWorld evaluation: Working under budget, time data, and political constraints. SAGE. Bamberger, M., Raimondo, E. and Vaessen, J. 2015. *Dealing with complexity by unpacking and reassembling elements of a complex program*. In Dealing with complexity in development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 127–143. SAGE.

Bamberger, M., Tarsilla, M. and Hesse-Biber, S. 2016. Why so many ‘rigorous’ evaluations fail to identify unintended consequences of development programs: How mixed methods can contribute. *Evaluation and Program Planning* 55: pp. 155–162.

Bamberger, M., Vaessen, J. and E. Raimondo. 2015. *Complexity in development evaluation: The framework of the book*. In Dealing with complexity in development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 1–25. SAGE.

Barnwal, P. 2017. Curbing leakage in public programs: Evidence from India’s direct benefit transfer policy. IGC Working Paper E-89111-INC-1. International Growth Centre.

Befani, B. 2012. *Models of causality and causal inference*. In Broadening the range of designs and methods for impact evaluations, E. Stern et al. (eds) Working Paper 38. DfID.

Befani, B. and Mayne, J. 2014. Process tracing and contribution analysis: A combined approach to impact evaluation. IDS Bulletin 45(6). Institute of Development Studies.

Berliner, D. and Dupuy, K. 2018. The promise and perils of data for anti-corruption efforts in international development work. U4 Brief 2018:7. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Bond. 2016. Choosing appropriate evaluation methods tool.

Building State Capability. 2018. PDIA toolkit: A DIY approach to solving complex problems. Center for International Development, Harvard University.

Burns, D. 2014. Assessing impact in dynamic and complex environments: Systemic action research and participatory systemic inquiry. Practice Paper Number 08. Centre for Development Impact.

Chayes, S. 2016b. The structure of corruption in Azerbaijan. In The structure of corruption: A systemic analysis using Eurasian cases. Carnegie Endowment for International Peace.

Cheeseman, N. and Peiffer, C. 2020. Why efforts to fight corruption hurt democracy: Lessons from a survey experiment in Nigeria. Working Paper 027. Anti-Corruption Evidence (ACE) Research Consortium.

Chelimsky, E. 2019. *Improving the match between sustainability questions and evaluation practice: Some reflections and a checklist*. In Evaluating Sustainability: Evaluative Support for Managing Processes in the Public Interest, Julnes, G. (ed.). New Directions for Evaluation 162: 69–86.

Council of Europe. 2013. Designing and implementing anti-corruption policies: Handbook

Dávid-Barrett, E. and Fazekas, M. 2020. Anti-corruption in aid-funded procurement: Is corruption reduced or merely displaced? World Development 132: 105000.

Dávid-Barrett, E. et al. 2020. Evaluating anti-corruption agencies: Learning from the Caribbean. Journal of Development Effectiveness 12(1): 74–88.

DfID. 2014. DfID Evaluation Strategy 2014–2019.

Espinosa, J. 2013. Moving towards gender-responsive evaluation? Practices and challenges in international-development evaluation. *Evaluation* 19(2): 171–182.

Feinstein, O. 2019. Dynamic evaluation for transformational change. In Evaluation for transformational change, van den Berg, R. D., Magro, C. and Mulder, S. S. (eds) 17–28. International Development Evaluation Association (IDEAS).

Fiala, N. and Premand, P. 2018. Social accountability and service delivery: Experimental evidence from Uganda. Policy Research Working Paper 8449. World Bank Group.

Fisman, R. and Golden, M. A. 2017a. Corruption: What everyone needs to know. Oxford University Press.

Fisman, R. and Golden, M. A. 2017b. How to fight corruption. *Science* 356(6340): 803–804.

Fletcher, G. 2015. Addressing gender in impact evaluation: What should be considered? ODI.

Foltz, J. D. and Opoku-Agyemang, K. A. 2015. Do higher salaries lower petty corruption? A policy experiment on West Africa's highways. International Growth Centre (IGC).

Garcia, M. 2011. Micro-methods in evaluating governance interventions. Evaluation Working Paper.

Gertler, P. J. et al. 2016. Impact evaluation in practice. Second edition. IBRD/World Bank.

Göbel, C. 2015. Anti-corruption in Taiwan: Process Tracing Report. ANTICORRP.

Hanna, R. et al. 2011. The effectiveness of anti-corruption policy: What has worked, what hasn't and what we don't know: A systematic review. EPPI-Centre.

- Hart, E. 2019. Guide to using corruption measurements and analysis tools for development programming. U4 Guide 2019:1. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Hart, R. A. 1992. Children's participation: From tokenism to citizenship. UNICEF.
- Heeks, R. 2011. Understanding success and failure of anti-corruption initiatives. U4 Brief 2011:2. Bergen: Norway: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Heywood, P. M. 2018. Combating corruption in the twenty-first century: New approaches. *Daedalus* 147(3): 83–97.
- ICAI. 2014. DFID's approach to anti-corruption and its impact on the poor. Report 37.
- IDEAS. 2019. Prague Declaration on Evaluation for Transformational Change.
- Independent Evaluation Group at the World Bank. 2016. Supporting transformational change for poverty reduction and shared prosperity: Lessons from World Bank Group experience.
- INTRAC (2017). Complex collection and analysis methodologies.
- Jackson, D. 2020. How change happens in anti-corruption: A map of policy perspectives. U4 Issue 2020:14. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Jackson, D. and Köbis, N. 2018. Anti-corruption through a social norms lens. U4 Issue 2018:7. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Jackson, D., Tobin, S. and Eggert, S. P. 2019. Capacity building for politicians in contexts of systemic corruption: Countering 'wasta' in Jordan. U4 Issue 2019:9. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Johnsøn, J. (2012). Theories of change in anti-corruption work: A tool for programme design and evaluation. U4 Issue 2012:6. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Johnsøn, J. and Søreide, T. 2013. Methods for learning what works and why in anticorruption: An introduction to evaluation methods for practitioners. U4 Issue 2013:8. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Johnsøn, J., Taxell, N. and Zaum, D. 2012. Mapping evidence gaps in anti-corruption: Assessing the state of the operationally relevant evidence on donors' actions and approaches to reducing corruption. U4 Issue 2012:7. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Johnsøn, J. et al. 2011. How to monitor and evaluation anti-corruption agencies: Guidelines for agencies, donors, and evaluators. U4 Issue 2011:8. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Johnston, M. 2014. Corruption, contention, and reform: The power of deep democratization. Cambridge University Press.

Johnston, M. and Johnsøn, J. 2014. Doing the wrong things for the right reasons? 'Do no harm' as a principle of reform. U4 Brief 2014:13. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Kane, R. et al. 2017. Contribution analysis in policy work: Assessing advocacy's influence. Center for Evaluation Innovation.

Khan, M., Roy, P. and Andreoni, A. 2019. Anti-corruption in adverse contexts: Strategies for improving implementation. Working Paper 013. Anti-Corruption Evidence (ACE) Research Consortium.

Kirya, M. n.d. Gender and anti-corruption programming. Research topic. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Klitgaard, R. (2015). Addressing corruption together. OECD. The Development Assistance Committee.

Kurer, O. 2015. Definitions of corruption. In Routledge handbook of political corruption, Heywood, P. M. (ed.) 30–41. Routledge.

Lemaître, S. 2019a. Corruption, évitement fiscal, blanchiment dans le secteur extractif: De l'art de jouer avec le droit. Presses Universitaires de Rennes.

Lemaître, S. 2019b. Illicit financial flows within the extractive industries sector: A glance at how legal requirements can be manipulated and diverted. *Crime, Law and Social Change* 71(1): 107–128.

Leouze, E., Areias, A. and Jackson, S. 2015. *The valuation of complex development interventions in the age of big data*. In Dealing with complexity in development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 221–250. SAGE.

Levy, B. 2014. Working with the grain: Integrating governance and growth in development strategies. Oxford University Press.

Marquette, H. and Peiffer, C. n.d. Corruption Functionality Framework. Global Integrity Anti-Corruption Evidence Research Programme.

Mason, P. 2020. Twenty years with anti-corruption. Part 8: Working with other parts of government ... when they don't want to work with you. U4 Practioner Experience Note 2020:8. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

McGee, R. and Gaventa, J. 2010. Review of impact and effectiveness of transparency and accountability initiatives: Synthesis report. Institute of Development Studies.

Mugellini, G., Villeneuve, J.-P. and Heide, M. 2021. Monitoring sustainable developments goals and the quest for high-quality indicators: Learning from a practical evaluation of data on corruption. *Sustainable Development* 29(6): 1257–1275.

Mungiu-Pippidi, A. 2017. Seven steps to evidence-based anti-corruption: A roadmap. Report 2017:10. EBA.

Mungiu-Pippidi, A. 2016. The quest for good governance: Learning from virtuous circles. *Journal of Democracy* 27(1): 95–109.

Mungiu-Pippidi, A. 2015. The quest for good governance: How societies develop control of corruption. Cambridge University Press.

Mungiu-Pippidi, A. and Dadašov, R. 2017. When do anti-corruption laws matter? The evidence on national integrity enabling contexts. *Crime, Law and Social Change* 68(4): 387–402.

Mæstad, O. and Mwisongo, A. 2017. Informal pay and the quality of health care: Lessons from Tanzania. U4 Brief 2007:9. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

ODI. 2012. RAPID Outcome Assessment.

ODI. 2009. Episode studies.

OECD. 2021. Applying evaluation criteria thoughtfully. OECD Publishing.

OECD. 2018. OECD DAC evaluation criteria: Summary of consultation responses.

OECD. n.d. Evaluation criteria.

OECD/DAC Network on Development Evaluation. 2019. Better criteria for better evaluation: Revised evaluation criteria definitions and principles for use.

Olken, B. A. 2005. Monitoring corruption: Evidence from a field experiment in Indonesia. Working Paper 11753. National Bureau of Economic Research.

Patton, M. 2020. Evaluation criteria for transformation: Webinar summary. Blue Marble Evaluation.

Patton, M. Q. 2012. Essentials of utilization-focused evaluation. SAGE Publishing.

Peiffer, C. et al. 2020. Lessons from reducing bribery in Uganda's health services. *Development Policy Review* 39(50): 721–739.

Persson, A., Rothstein, B. and Teorell, J. 2019. Getting the basic nature of systemic corruption right: A reply to Marquette and Peiffer. *Governance* 32(4): 799–810.

Raimondo, E. and Bamberger, M. 2015. *Gender equality in development evaluation: The intersection of complexities*. In Dealing with complexity in

development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 273–292. SAGE.

Raimondo, E., Vaessen, J. and Bamberger, M. 2015. *Towards more complexity-responsive evaluations: Overview and challenges*. In *Dealing with complexity in development evaluation: A practical approach*, Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 26–47. SAGE.

Robillard, S. and Robillard, L. 2018. *When not to call a spade a spade: The importance of quiet anti-corruption initiatives*. Blog: Corruption in Fragile States.

Rocha Menocal, A. et al. 2015. *Why corruption matters: Understanding causes, effects and how to address them*. Evidence Paper. DfID

Roelen, K. and Devereux, S. 2014. *Evaluating outside the box: Mixing methods in analysing social protection programmes*. Practice Paper Number 06. Center for Development Impact.

Rothstein, B. 2011. *Anti-corruption: The indirect ‘big bang’ approach*. Review of International Political Economy 18(2): 228–250.

Rothstein, B. and Tannenbergh, M. 2015. *Making development work: The quality of government approach*. Report 2015:07. Expert Group for Aid Studies (EBA).

Rothstein, B. and Varraich, A. 2017. *Making sense of corruption*. Cambridge University Press.

Scharbatke-Church, C. 2017. *Research methodology for identifying social norms that catalyze corruption*. CDA.

Scharbatke-Church, C. 2016. *What dynamics drive police and judicial officers to engage in corruption*. Blog: Corruption in Fragile States.

Scharbatke-Church, C. and D. Chigas .2016. *Taking the blinders off: Questioning how development assistance is used to combat corruption*. Occasional Paper. Institute for Human Security.

Scharbatke-Church, C. and Kothari, D. 2021. *Something old, something new, something borrowed & something blue*. Working Paper. CJL.

- Scharbatke-Church, C. with Barnard-Webster, K. 2017. Understanding corruption in criminal justice as a robust and resilient system: An analysis process using systems thinking tools. CDA Collaborative Learning Projects.
- Scharbatke-Church, C., Barnard-Webster, K. and Woodrow, P. 2017. Collective action against corruption in the criminal justice system. CDA Collaborative Learning Projects.
- Schütte, S. A. 2019. Why fix KPK when it is not broken? Jakarta Post.
- Schütte, S. A. 2017. Bespoke monitoring and evaluation of anti-corruption agencies. U4 Brief 2017:2. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Sida. 2020. Sida's evaluation handbook: Guidelines and manual for conducting evaluations at Sida.
- Smith, D. J. 2015. *The contradictions of corruption in Nigeria*. In Routledge handbook of political corruption, Heywood, P. M. (ed.) 56–66.
- Spector, B. I., Winbourne, S. and Dininio, P. 2015. Practitioner's guide for anticorruption programming. USAID.
- Stern, E. et al. 2012. Broadening the range of designs and methods for impact evaluations. Working Paper 38. DfID.
- Stevens, A. 2016. Configurations of corruption: A cross-national qualitative comparative analysis of levels of perceived corruption. *International Journal of Comparative Sociology* 57(4): 183–206.
- Taylor, M. M. 2018. Getting to accountability: A framework for planning & implementing anticorruption strategies. *Daedalus* 147(3): 63–82.
- Transparency International. Corruption Perceptions Index 2021.
- Trapnell, S. E. 2015. User's guide – measuring corruption and anti-corruption. UNDP.

Trapnell, S., Jenkins, M. and Chêne, M. 2017. Monitoring corruption and anti-corruption in the Sustainability Development Goals: A resource guide. Transparency International.

U4. 2017. Managing a hostile court environment: Common challenges and recommendations. Notes from the Corruption Hunter’s meeting, 11-12 November 2017.

U4 Information. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

U4. n.d. Glossary. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Uberti, L. J. 2020. Identifying feasible, high-impact anti-corruption interventions: The case of Albania. U4 Issue 2020:9. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

UNEG. 2016. Norms and standards for evaluation.

UNODC. 2017. Evaluation Handbook: Guidance for designing, conducting and using independent evaluation.

Vaessen, J., Raimondo, E. and Bamberger, M. 2015. *Impact evaluation approaches and complexity.* In Dealing with complexity in development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 61–87. SAGE.

Vaillant et al. 2020. Evaluation of Norway’s anti-corruption efforts as part of its development policy and assistance. Norad.

van de Braak, S., Choenni, S. and Bamberger, M. 2015. *Emergent technologies and creative use of multiple sources of information.* In Dealing with complexity in development evaluation: A practical approach. Bamberger, M., Vaessen, J. and Raimondo, E. (eds) 181–199. SAGE.

Wathne, C. 2021. Understanding corruption and how to curb it: A synthesis of latest thinking. U4 Issue 2021:3. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Wedel, J. I. n.d. The Mapping Method: A guide to charting corruption and influence processes. Global Integrity Anti-Corruption Evidence Research Programme.

White, H. 2019. The twenty-first century experimenting society: The four waves of the evidence revolution. Palgrave Communications 5(47).

Wilson, B. M. and Fernández, E. V. 2015. Process-tracing report on Costa Rica. European Research Centre for Anti-Corruption and State-Building.

Winbourne, S. and Spector, B. I. 2014. Analysis of USAID anticorruption programming worldwide (2007–2013): Final Report. USAID.

Woolcock, M. 2013. Using case studies to explore the external validity of ‘complex’ development interventions. Evaluation 19(3): 229–248.

World Bank. 2020. Enhancing government effectiveness and transparency: The fight against corruption.

World Bank/IDB. 2016. Impact evaluation, in practice. Second edition.