Bespoke monitoring and evaluation of anti-corruption agencies

Photo: patrick_q on flickr.com

Anti-corruption agencies (ACAs) are often considered a last resort against corruption and are expected to solve a problem that other institutions have failed to address effectively or may even be part of. When national anti-corruption strategies yield no result, and a country's corruption rankings do not improve, ACAs often take the brunt of the criticism. External institutional assessments can help to pinpoint successes and shortcomings and often shape public perceptions, but they cannot replace the consistent collection of data on progress by ACAs themselves. ACAs can proactively manage expectations as part of their public outreach and show their worth by clearly stating their objectives at the output and outcome levels and by collecting and analysing information on a raft of indicators. Measuring the completion and outcome of activities by the ACA can also provide important lessons on what works and what does not, allowing the agency to adjust its approaches.

Anti-Corruption Resource Centre

www.U4.no

Sofie Arjon Schütte

Senior advisor (U4/CMI)

In the last two decades, more than 30 countries have established new anti-corruption agencies (Recanatini 2011). Most of these agencies have been set up in developing countries, and in most cases donors have provided considerable technical assistance and financial support. ACAs are often seen as a last resort to reduce corruption, but these unrealistic expectations have quickly given way to perceptions of failure when the institutions have not

delivered the hoped-for results. Corruption has not been eradicated, and in some countries corruption perception indicators may even have worsened since the ACA's inception. This Brief does not seek to defend the disappointing performance of some ACAs or to reinstall them as the ultimate solution to corruption. Rather, it provides reflection and guidance on how to fairly measure the results achieved by ACAs and how to learn from and act on these findings.

This Brief draws largely on the 2011 U4 Issue *How to Monitor and Evaluate Anti-Corruption Agencies: Guidelines for Agencies, Donors, and Evaluators* (Johnsøn et al. 2011).[1] That paper emphasised the need for an informed debate – not on whether ACAs are good or bad per se, but on *why, to what extent,* and *in which contexts* they do or do not deliver results. To inform an evidence-based debate, assessments of ACAs should not only function as organisational audits, using a checklist to assess whether the agency has all its prescribed powers and organisational structures and exhibits sufficient independence. This normative approach, measuring an ACA against predefined standards, allows for cross-country and cross-agency comparisons; it may even give the ACA management some political leverage towards government, enabling it to argue for more autonomy or resources. Another dimension is often overlooked, but is more immediately useful in helping ACAs improve their performance relative to their mandate, powers, and resources. It consists of regular monitoring and evaluations (M&E) of the agency's interventions, using bespoke indicators to measure selected outputs and outcomes. The results of such M&E serve as valuable institutional feedback and learning tools.

Who is responsible for outputs, outcomes, and impact?

Public expectations are often high for ACAs, but the agencies do not always receive the resources, political space, or time necessary to live up to these expectations. Expectations of direct causal effects from ACA interventions at the outcome and impact levels may be unrealistic. In most areas of anti-corruption work, researchers have not yet even solved overall causality problems. For instance, we

understand that there is a correlation between transparency and lower levels of corruption. But we do not know with certainty whether transparency leads to less corruption, or whether transparency measures are simply more likely to be introduced in less corrupt environments. If we encourage ACAs to promote transparency measures but we cannot show a direct causal effect on reduction of corruption, that does not mean that the ACAs are necessarily to blame.

Generally, the first question to ask is whether a change is observable. The second question is whether the observed change can be attributed to the intervention. Questions on causality, impact, and attribution are methodologically challenging. Answering them involves building a counterfactual scenario in which no intervention has been administered.

When it is not possible to establish attribution, a contribution analysis can be conducted instead. Rather than trying to assess the proportion of change that results from the intervention (attribution), a contribution analysis aims to make a more limited assertion, namely, that the evaluated intervention is or is not one of the causes of observed change. Contributions may be ranked (in terms of influence relative to other factors), but not quantified. The analysis takes a step-by-step approach, building a chain of logical arguments while taking into account alternative explanations and relevant external factors.

Designing a strong results-based management framework, theory of change, or logical framework can help identify what results the organisation can be held accountable for, given its resources and constraints. If resources are available, the ACA should produce the desired outputs and can be held accountable for them. At the outcome level, however, the results are no longer completely within the ACA's control. If the intervention logic holds, the production of outputs should lead to the desired outcomes, but external factors such as changes in funding or the political climate may interfere. The outcome level is thus a grey area for accountability. ACAs are responsible, but only to a certain extent.

It is useful to distinguish between the spheres of ACA concern and responsibility (figure 1). The immediate responsibility of the ACA is for its output – for example, the successful investigation and prosecution of corruption cases, or raising awareness
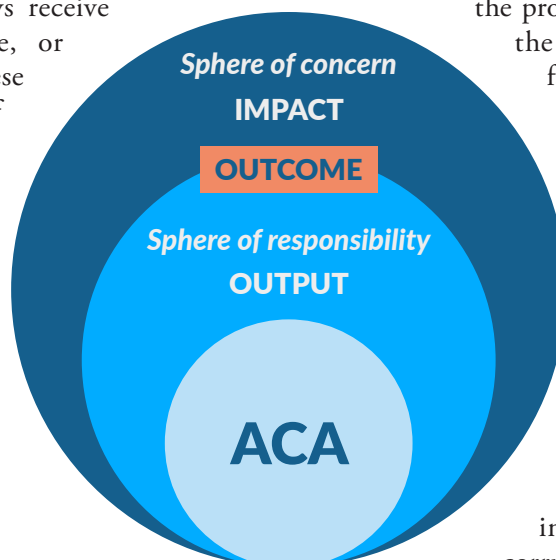
Figure 1: Spheres of concern and responsibility of an ACA

of corruption. If the results chain holds, the output will contribute to the desired outcome or impact, but other factors may interfere with that. For example, a national anti-corruption strategy is clearly within the sphere of concern of an ACA, but not all elements of this strategy may be within the sphere of its responsibility. All stakeholders in a national anti-corruption strategy can contribute to the overall objective, but no single stakeholder can be held directly accountable for delivering on the entire strategy, or claim its success at the highest level.

Why develop a bespoke monitoring and evaluation system for an anti-corruption agency?

External assessment can provide important benchmarks for an ACA, but it cannot replace continuous monitoring and regular evaluations by the ACA itself. The development of bespoke objectives and indicators enables the ACA to learn whether a specific intervention has yielded the desired outputs and outcomes, and if not, why not. This provides the necessary detail for institutional learning. Ideally, a dedicated monitoring and evaluation unit will be set up within each ACA to design and coordinate bespoke data collection and conduct evaluations. This, of course, requires adequate allocation of resources to and within the ACA.

The U4 Issue (Johnsøn et al. 2011) recommends that evaluation processes should have built-in mechanisms for dissemination of the report and for feedback. This applies equally, if not more so, to monitoring. When evaluations are done right – and when they are not mere accountability exercises for outsiders – they provide important information for future policy and programme development. Here again, a participatory approach is vital to ensure a transparent process. Formal mechanisms for feedback could include scheduling review processes, peer reviews, seminars and workshops, etc. As observed by the OECD, "In order to be effective, the feedback process requires staff and budget resources as well as support by senior management and the other actors involved" (2010, 10). For this to happen, the U4 Issue recommends that feedback mechanisms be planned and budgeted as part of the monitoring evaluation process. A formal management response and follow-up mechanism – for example, through an action plan – should be developed to systematise implementation of the recommendations. The M&E unit, together with senior management, should ensure that all agreed follow-up actions are tracked to ensure accountability for their implementation. Later evaluations should review how management and staff responded to the recommendations, whether they were implemented,

and whether the desired effect was produced (Johnsøn et al. 2011, 57).

Perhaps the biggest challenge is to promote a culture within the ACA that sees monitoring and evaluation as a useful learning tool rather than as an unwelcome duty or even as threatening scrutiny.

How to construct indicators

Indicators are an essential instrument in M&E systems but are often misunderstood, poorly constructed, and misapplied. An indicator can be defined as a measure tracked systematically over time that indicates positive, negative, or no change with respect to progress towards a stated objective. The U4 Issue recommends that managers or policy makers examine the combined evidence from a group of indicators to evaluate whether the intervention is having positive effects. They should not attempt to measure any outcome or impact using only one indicator. Indicators are normally derived from the impact, outcomes, and outputs defined in advance as desired results. It is therefore important to establish those clearly and to ensure that they follow a clear logic or theory of change (Johnsøn et al. 2011, 47).

Disaggregation of indicators, to the extent possible, can help capture differences in, for example, types of corruption, corruption by sector (public, private, police, customs, etc.), gender, locality, and methods of reporting corruption (e-mail, letter, personal, etc.).

U4 encourages the use of mixed methods, both in the traditional sense of mixing quantitative and qualitative indicators (e.g., quantitative statistics and qualitative case studies) and also in an extended sense – by mixing indicators that are perception-based, or proxies, with "harder" indicators that directly measure corruption. For a list of the advantages of mixed methods, see the U4 Issue (Johnsøn et al. 2011, 40–41).

Unfortunately, purely empirical "objective" anti-corruption indicators based on hard facts are difficult to find. As a result, the use of proxy indicators is common in the social sciences. Many proxy indicators, if used correctly, can yield very good approximations to reality. However, one must be aware of the nature of an indicator when interpreting it. "The number of corruption cases brought to trial," for example, should not be seen as a proxy indicator for corruption levels in the country, because an increase in cases brought to trial could indicate a higher incidence of corruption, an increased level of confidence in the courts, or both. This indicator is, rather, a proxy for the efforts of investigators, prosecutors, and the judiciary (Schütte and Butt 2013).

Why is monitoring and evaluation important to ACAs' work and existence?

Transparency: Visibility and outreach are important in attracting public support. Public ignorance about the existence and functioning of an ACA lays the conditions for its marginalisation or gradual death. M&E generates written reports and concrete performance figures that contribute to transparency and visibility.

Accountability: ACAs are publicly funded bodies. Therefore, they need to report on their activities, capacity problems, and results to those who fund their activities: taxpayers and donors. M&E provides reliable information on performance and makes it possible to track progress more easily and systematically.

Institutional memory: M&E allows the ACA to develop an in-house memory about the different phases of institutionalisation.

Learning: M&E provides evidence for questioning and testing assumptions, and for adjusting strategies, policies, and practices. It offers a basis and a process for self-reflection.

Improving policy: M&E frameworks can give the heads of ACAs and governments indications of whether a policy option is working as intended by detecting operating risks and problems. Where do the problems originate? How do they affect the agency's performance? What capacities/resources are available to reduce those risks and problems, and can those be strengthened?

Better performance: All of the above can contribute to better performance of ACAs in fighting corruption. It also provides ACAs with a more robust basis for raising funds and influencing policy.

Adapted from Johnsøn et al. (2011, 16).

Similarly, when dealing with perception-based indicators, one should remember that they reflect people's subjective opinions. These might be influenced by, for example, increased media reporting on an ACA's investigation of corruption cases, which may or may not reflect an increase in corruption itself. In sum, triangulation of different kinds of indicators and different sources of verification is essential to strengthen the validity and reliability of the findings.

Under the United Nations 2030 Agenda for Sustainable Development, governments agreed to identify country-level indicators to measure their progress towards the Sustainable Development Goals and targets, including SDG target 16.5, which calls on countries to "substantially reduce corruption and bribery in all their forms." This has reignited the discussion on how to measure a complex phenomenon such as corruption using a handful of indicators, if not a single indicator. The indicator proposed for target 16.5 – the percentage of persons who paid a bribe to a public official, or were asked for a bribe by a public official, during the last 12 months – is a useful experience-based indicator for bribery and extortion. But it will not generate information about the types and value of bribes, nor about types of corruption other than bribery. Nonetheless, the surveying of just this one indicator by national statistical offices can provide an immense opportunity to ACAs. It can offer an entry point for adding more indicators, such as on bribery experiences in specific sectors and on other types of corruption.

Finally, an ACA can conduct its own data collection beyond the data generated as part of its operations. An example is the annual survey conducted by ICAC in Hong Kong, which measures, among other things, public attitudes towards corruption and the work of the ICAC and reasons for these attitudes. The surveys are conducted by independent research firms (ICAC 2016).

Existing assessment resources and tools

Since the publication of the U4 Issue in 2011, new guidelines and assessment tools have been prepared to help ACAs develop bespoke monitoring and evaluation. The two most prominent guides, each of which has been applied in several countries, are briefly described below.

Anti-corruption agencies strengthening initiative

In 2015, Transparency International published a guide to a participatory assessment methodology for anti-corruption agencies. The international NGO plans to carry out regular independent assessments "to capture internal and external factors affecting the ACA as well as getting a sense of the ACA's reputation and actual performance" (León 2015, 10). A total of 50 indicators with three possible scores (high, moderate, and low) are measured across seven dimensions of common relevance for most ACAs (Quah 2015):

- Legal independence and status
- Financial and human resources
- Detection and investigation function
- Prevention, education, and outreach functions
- Cooperation with other organisations
- Accountability and oversight
- Public perceptions of the ACA's performance

The guidelines contain fixed questions and values around the design and effectiveness of ACAs; accordingly, these

assessments can be used to compare and rank the agencies. Given that ACAs do not operate in a vacuum, the assessment takes into account external enabling and constraining factors as well as performance of key functions. It is not primarily a tool that ACAs can use to measure the effectiveness of their organisational strategies and specific interventions.

By early 2017, Transparency International had conducted such assessments in Bangladesh, Bhutan, Indonesia, Maldives, Mongolia, Sri Lanka, Pakistan, and Taiwan. To date, only the assessments from Bhutan and Bangladesh are publicly available (León 2015; Aminuzzaman, Akram, and Islam 2016). It is too early to say what impact these assessments have had and whether they have brought about any changes within the ACAs or the environments in which they operate.

Practitioner's Guide: Capacity assessment of anti-corruption agencies

The *Practitioner's Guide* developed by the United Nations Development Programme offers instructions for critically assessing the capacity of an ACA at the individual, organisational, and environmental levels (UNDP 2011). In terms of its purpose, a capacity assessment is quite distinct from an evaluation, which is typically more *performance*-oriented than *needs*-oriented. But many of the questions used to assess capacity could also be used to construct output and outcome objectives and respective indicators for monitoring and evaluation. For example, a question from the *Practitioner's Guide* on internal cooperation and synergies between investigative and preventive functions could be turned into objectives and indicators as follows:

Question from *Practitioner's Guide*: When systematic patterns of corruption are diagnosed by the investigation team or the prevention team, are these findings shared among both teams?

This question could be translated into the following output objective: *The ACA's investigative work and preventive work reinforce each other through knowledge and intelligence sharing and allow for resource prioritisation.*

This objective in turn could be measured by the following output indicators:

- Number of coordination meetings between investigative and preventive staff
- Number of files transmitted between investigative and preventive staff
- Rotation or co-placement of staff across teams or joint teams

This would not tell us anything about the effect, or desired outcome, that this cooperation might have. Such an outcome could be stated as: *Government institution (one or more) that has previously been under investigation by the ACA conducts a systems review and adopts preventive measures.*[2]

Indicators for this outcome objective could be:

- Number of ACA investigations that lead to follow-up preventive interventions (e.g., specific control tools) at previously investigated institutions.
- Proportion of recommendations from systems review that are adopted by the designated institutions (here a qualitative weighing of the recommendations might also be useful).

For most ACA interventions, the ultimate impact objective presumably would be that *corruption is sustainably reduced in a particular institution/area of intervention for which investigative and preventive teams have joined forces.*

This could be measured, at least in part, by the following impact indicators:

- Perception and incidences of (specific types of) corruption experienced have declined in number (based on survey).
- Clients and customers of that particular institution/sector state higher satisfaction with services (e.g., as regards speed of services, accessibility, transparency, quality).

Note that a change in these impact indicators cannot necessarily be attributed to greater collaboration between the ACA's different teams, or even to the work of the ACA at all. External factors beyond the ACA's influence, such as new leadership in the institution under investigation, may also contribute to the change. The further an objective is beyond the control of the ACA, the more difficult it gets to ascertain causality. In the case here, indicators for causality could be bolstered by interviews and focus-group discussions of key personnel in the institution/ sector that has been under ACA scrutiny.

While the development of a functioning M&E system is the responsibility of the ACA, external evaluators from civil society, academia, and donor agencies can play a constructive role by providing advice, quality assurance, and critique. They may provide external validation of the assessment methodology and/or carry out additional external assessments to broaden the perspective, for example by looking at the enabling environment (as does the Transparency International assessment mentioned above).

For the purpose of internal monitoring and evaluation, we offer the following recommendations to ACAs and their governments, and in particular to supporting donors:

- Proactively manage public expectations by clearly stating the ACA's objectives at the output and outcome levels.
- Allocate sufficient resources, including special M&E expertise, to generate in-house data and retrieve and analyse relevant external data.
- Use a mix of various types of indicators (perception, experience, proxy, milestones).
- Remember that learning is integral to M&E. There are strong pressures on ACAs to show results, but this should not overshadow the need to learn what works and why.

References

Aminuzzaman, S., S. M. Akram, and S. L. Islam. 2016. *Anti-Corruption Agency Strengthening Initiative: Assessment of the Bangladesh Anti-Corruption Commission 2016*. Transparency International Bangladesh. https://goo.gl/d6P0t7.

ICAC (Independent Commission against Corruption). 2016. Annual Survey. Hong Kong. https://goo.gl/RTlTHZ.

Johnsøn, J., H. Hechler, L. De Sousa, and H. Mathisen. 2011. *How to Monitor and Evaluate Anti-Corruption Agencies: Guidelines for Agencies, Donors, and Evaluators*. U4 Issue 2011:8. Bergen, Norway: Chr. Michelsen Institute. https://goo.gl/2Et0HL.

León, J. R. 2015. *Anti-Corruption Agency Strengthening Initiative: Assessment of the Bhutan Anti-Corruption Commission 2015*. Transparency International. https://goo.gl/hZDCbz.

OECD (Organisation for Economic Co-operation and Development). 2010. *Evaluating Development Co-operation: Summary of Key Norms and Standards*. 2nd ed. Paris: OECD DAC Network on Development Cooperation. https://goo.gl/hUJxmP.

Quah, J. 2015. *Anti-Corruption Agencies Strengthening Initiative: Research Implementation Guide.* Berlin: Transparency International. https://goo.gl/gs4RMI.

Recanatini, F. 2011. "Anti-Corruption Authorities: An Effective Tool to Curb Corruption?" In *International Handbook on the Economics of Corruption*, vol. 2, edited by S. Rose-Ackerman and T. Søreide, 528–69. Cheltenham, UK: Edward Elgar.

Schütte, S., and S. Butt. 2013. *The Indonesian Court for Corruption Crimes: Circumventing Judicial Impropriety?* U4 Brief 15:2013. Bergen, Norway: Chr. Michelsen Institute. https://goo.gl/Ob0aut.

UNDP (United Nations Development Programme). 2011. *Practitioner's Guide: Capacity Assessment of Anti-Corruption Agencies.* New York.

Endnotes

1. The author would like to thank Jesper Johnsøn, Hannes Hechler, Luís De Sousa, and Harald Mathisen, the authors of the original U4 Issue, for permission to use and adapt sections of their text. Thanks also to Kirsty Cunningham, Fredrik Eriksson, Rukshana Nanayakkara, and Marijana Trivunovic for their useful comments on a draft of this Brief.

2. An additional outcome objective and related indicators could be added to cover analysis by the prevention team leading to focused investigative action.