



GLOBAL INTEGRITY
Innovations for Transparency & Accountability



*Empowered lives.
Resilient nations.*



United Nations Development Programme

User's Guide to
**MEASURING CORRUPTION
AND ANTI-CORRUPTION**

AUTHOR

Stephanie E. Trapnell

MANAGING EDITORS

Hazel Feigenblatt, Global Integrity
Diana Torres, UNDP

EDITORS

Anga R. Timilsina, UNDP

Acknowledgments

User's Guide to Measuring Corruption and Anti-corruption was commissioned by the UNDP Global Anti-corruption Initiative (GAIN). We would like to thank Global Integrity for its crucial assistance with preparing this Guide. We are indebted to the following practitioners for serving as the peer reviewers: Constanze von Soehnen (UNODC), Craig Fagan (TI), Estelle Fach (UN-REDD programme); Francesco Checchi (formerly UNDP and now with UNODC), Fredrik Galtung (Integrity Action), Jesper Johnson (U4 Resource Centre), Diego Antoni (UNDP), Johannes Ferguson (GIZ), Phil Mason (DFID), Serge Kapto (UNDP), Oskar Karneback (OECD), Elodie Beth (UNDP), and Njoya Tikum (UNDP). We are also grateful to Richard E. Messick (Independent Expert) and Aida Arutyunova, Israel Marañon and Sabrina Hoque (UNDP GAIN Team) for their valuable contributions. We would also like to thank UNDP Malaysia for providing support to design and print this Guide.

Finally, this Guide could not have been written without cooperation from our interviewees (listed in annex). We are grateful for their insights and honesty when speaking with us. We would also like to thank the Department of Foreign Affairs and Trade (DFAT) of Australia for its generous funding support to produce this Guide.

USER'S GUIDE TO MEASURING CORRUPTION AND ANTI-CORRUPTION.

Copyright © 2015 by UNDP. All rights reserved. For information regarding the appropriate use of this document, contact UNDP Global Anti-corruption Initiative (GAIN) at gain@undp.org.

Cover and interior design by Torquee Advertising Sdn Bhd

Contact:

United Nations Development Programme

One United Nations Plaza

New York, NY, 10017

Tel: +1 212 906 5081

Email: gain@undp.org

<http://www.undp.org/governance>

<http://www.anti-corruption.org>

Contents

Acknowledgments	5
Abbreviations	6
Tables and Figures	7
Foreword	8
Introduction	9
Practitioner's Quick Reference	9
Chapter 1. Data, Methods, and Measurements of Corruption/Anti-corruption	11
Macro vs. Micro: Composite Indicators and Policy-relevant Data	13
Approaches to Measurement: Data and Methods	14
Types of Data	14
Common Methods	17
Measurements of Corruption	20
Measurements of Anti-corruption: Transparency, Accountability, Integrity	21
Transparency	22
Accountability	25
Integrity	29
Chapter 2. Measurement and the M&E Cycle	31
Logic Models	32
Theories of Change	34
Measurement Planning	41
How Can It Be Measured? Finding the Data	43
Methodologies: Considerations and Constraints	47
Measurement Design Considerations	51
Stakeholders	51
De Jure vs. De Facto Measurements	53
Quality Control	54
Sustainability	57
Constraints	57
Chapter 3. Evaluation and Impact Assessment	59
Experimental Impact Evaluations	60
Theory-based Evaluation	63
Most Significant Change	64
Outcome Mapping	65
Process Tracing	66
Other Evaluation Methods	67
Approaches to Measuring Effectiveness at the Project/Sector Level	69
Advocacy Campaigns	69
Mitigating Risks in Sectors	70
Empowerment and Capacity Building	75
Approaches to Measuring Effectiveness at the Institutional/Organizational Level	77
Anti-corruption Bodies	77
Institutional Integrity	79
Approaches to Measuring Effectiveness at the National Level	81
National Anti-corruption Strategies/Policies	81
International and Regional Conventions/Treaties	83

Contents

Chapter 4. Lessons Learned from the Trenches (Practitioner Voices)	88
Data Should Not Be Produced in a Vacuum	88
Data is Not the End of the Story	90
There is Still a Need for Global, Comparative Data	94
Recognize the Complementarity (Not Comparability) of Different Kinds of Measurements	95
“Local Data”: Effective at Guiding Reforms, Difficult to Sustain Over Time	97
Be Careful of Negative Benchmarking and Gaming	100
Ensure that Data is Credible, Without Complex Statistical Techniques	100
In Turn, Data Can Build Credibility in Reform Efforts	101
Always Ask – What is the Purpose of the Data Set? What is it Designed to Do?	102
Be Precise, Even with Fuzzy Concepts	104
Beware of Easy Answers	105
Chapter 5. Good Practices for Accountability	107
Spotlight on Good Practices	107
Complementarity: Context Matters. Purpose Matters	110
Knowledge of the Commons	111
Partnership and Sustainability	112
References	113
Annex 1 : List of Interviewees	118
Annex 2 : Datasets	119
Annex 3 : Experimental and Quasi-experimental Methods	121
Annex 4 : Summary of Selected Studies Using “Objective” Indicators of Corruption	122
Annex 5 : Institutional Transparency Indicators	123
Annex 6 : Institutional Accountability Indicators	124

Abbreviations

ACA	Anti-corruption Agency
ACN	Anti-corruption Network
ACRC	Korea Anti-corruption & Civil Rights Commission
AIA	Anti-corruption Initiative Assessment
ALACs	Advocacy and Legal Advice Centres
APRM	African Peer Review Mechanism
BEEPS	Business Environment and Enterprise Survey
BPP	Bureau of Public Procurement
CBOs	Community Based Organizations
CEPEJ	European Commission for the Efficiency of Justice
CPA	Commission on Audit's Citizen Participatory Audit
CPI	Corruptions Perceptions Index
CPIA	Country Performance and Institutional Assessment
DFAT	Department of Foreign Affairs and Trade
DFID	Department for International Development
EDI	Environmental Democracy Index
EITI	Extractive Industries Transparency Initiative
FATF	Financial Action Task Force
GAC	Governance and Corruption Diagnostics
GAIN	Global Anti-corruption Initiative
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
GKA	Global Key Awards
GRECO	Council of Europe Group of States against Corruption
IA	Integrity Assessment
IBP	International Budget Partnership
IDR	Integrity Development Review
IMF	International Monetary Fund
IP	Integrity Pact
IPAR	Institute for Policy Analysis and Research
IRM	Independent Reporting Mechanism
MAPS	Methodology for Assessing Procurement Systems
MESICIC	Mechanism for Follow-up on the Implementation of the Inter-American Convention against Corruption
MGNREGS	Mahatma Gandhi National Rural Employment Guarantee Scheme
NACC	National Anti-corruption Commission
NIS	National Integrity Systems
OAS	Organization of American States
OBI	Open Budget Index
OECD	Organization for Economic Cooperation and Development
OGP	Open Government Partnership
PACDE	Global Programme on Anti-corruption for Development Effectiveness
PAPI	Public Administration Performance Index
PEFA	Public Expenditure and Financial Accountability
PETS	Public Expenditure Tracking Surveys
PM&E	Participatory Monitoring and Evaluation
PPDC	Public and Private Development Centre
RAAG	RTI Assessment and Analysis Group
RTH	Right to Hearing
RTI	Right to Information
SHaSA	Strategy for the Harmonization of Statistics in Africa
SMART	Specific, Measurable, Attainable, Relevant, Time-bound
TADAT	Tax Administration Diagnostic Assessment Tool
TI	Transparency International
UNCAC	United Nations Convention Against Corruption
UNDP	United Nations Development Programme
UNICEF	United Nations Children's Fund
UNODC	United Nations Office of Drugs and Crime
UN-REDD	United Nations Programme on Reducing Emissions from Deforestation and Forest Degradation
WJP	World Justice Project

Tables and Figures

Table 1: Drawbacks of Aggregate Indicators	12
Table 2: Selected Measurements of Corruption	20
Table 3: Selected Examples of RTI Measurements	22, 23
Table 4: Selected Examples of Public Sector Transparency Measurements	24
Table 5: Selected Examples of Measurements of Institutional or Horizontal Accountability	26
Table 6: Selected Examples of Vertical Accountability Measurements	28
Table 7: Selected Examples of Rule of Law Measurements	29
Table 8: Selected Examples of Integrity Measurements	30
Table 9: Possible Indicators and Data Sources for Targeting Corruption	38
Table 10: Possible Indicators and Data Sources for Improving Government Transparency	39
Table 11: Possible Indicators and Data Sources for Enhancing Government Accountability	40
Table 12: Selected Pro-poor and Gender Sensitive Indicators for Corruption	44
Table 13: Selected Methodologies, with Factors to Consider when Adapting to Different Contexts	48, 49, 50
Table 14: Measurements of External Outcomes in Advocacy Efforts	70
Table 15: Key Tools to Identify, Track and Measure Corruption Risks and Corruption in the Health Sector	73
Table 16: Anti-Corruption Monitoring Mechanisms for the Forestry Sector	74, 75
Table 17: Selected Indicators for Performance Monitoring and Outcome Evaluation of Anti-corruption Bodies	78
Table 18: Country Examples of Indicators from Anti-corruption Bodies	79
Table 19: Country Examples of Institutional-level Assessment Tools	80
Table 20: Number of Countries (40 Total) with Specific Themes Emerging from Recommendations and Observations Generated from GRECO Reporting Mechanisms	86
Figure 1: Global vs Local: What Questions Can the Data Answer?	14
Figure 2: A Continuum of Data Types Found in Selected Methods and Datasets	15
Figure 3: Tips When Assessing the Impact of Anti-corruption Work	20
Figure 4: Relationship Between Measurement/Data and M&E Processes	32
Figure 5: Results Chain for Corruption and Anti-corruption Interventions	33
Figure 6: Five-step Approach to Building a Theory of Change in Anti-corruption Interventions	35
Figure 7: Theory of Change Analysis for Anti-corruption Enforcement Work	35
Figure 8a: Logic Model with Theory of Change	36
Figure 8b: Logic Model with Theory of Change (An Example from UNDP's Global Anti- Corruption Initiative)	37
Figure 9: Four Strategic Questions to Guide Measurement Planning	42
Figure 10: Finding Relevant Data to Meet Project Needs	43
Figure 11: Disaggregation of Sample Corruption Sector	45
Figure 12: Considerations When Using Existing Data	46
Figure 13: Summary of the Evidence from a Systematic Review of Anti-corruption Interventions	60
Figure 14: Risk Assessment as Combination of Likelihood and Impact	71
Figure 15: Likelihood and Consequences of Risk Factors: Risk Cataloging	71
Figure 16: Capacity Assessment of Anti-corruption Agencies (UNDP Methodology)	78
Figure 17: OGP Commitments with Significant or Complete Progress	84
Figure 18: 3-dimensional Mapping of the Purpose for Measurement Initiatives	102
Figure 19: Governance – Corruption Continuum	104

Foreword

Corruption and anti-corruption measurements are increasingly recognized as essential tools for analyzing corruption trends and for monitoring the results of efforts to curb corruption, as evident from the rapid growth in the production and use of governance and anti-corruption indicators over the past two decades.

With easier access to more comprehensive and more rigorous corruption data, national and international actors alike have been able to design more effective policies and normative instruments to address a wide range of drivers of corruption, such as inequities in service delivery, poverty and exclusion, as well as loopholes in the international financing system, which fuel money laundering, tax evasion and international bribery.

Notwithstanding these significant advances, we continue to face serious challenges when trying to quantify the exact costs and consequences of corruption at global, regional, national and local levels. First and foremost, corrupt practices are generally hidden, making them difficult to identify and account for. Furthermore, the fact that corruption takes various forms – ranging from embezzlement, fraud, nepotism, bribery, extortion and money laundering – makes it impossible to capture corrupt practices in a single indicator. Corruption is also often underreported given victims may fear retaliation or may, to some extent, share responsibility for the crime.

Today, in spite of the vast amounts of resources invested in curbing corruption, we often find ourselves unable to fully document and report on the results achieved. Worse still, limited evidence of the effectiveness of anti-corruption programmes may translate mistakenly into unfavorable reviews of recently established anti-corruption agencies and national anti-corruption strategies.

This Guide builds on a previous edition published in 2008 by UNDP with Global Integrity, **“A User’s Guide to Measuring Corruption”**. The new edition aims to tackle this growing challenge by proposing methods and tools to capture the progress and impact of anti-corruption programmes. Recognizing that there is no single formula or one-size-fits-all approach to unravel the complexity of these issues, we have refrained from providing prescriptive guidelines. Rather, a range of tools is presented as well as general principles to consider when designing sound anti-corruption assessments or evaluations.

As the world gears up for the implementation of the Post-2015 Development Agenda, we hope that this Guide will prove useful for tracking the implementation of Goal 16 on “Promoting peaceful and inclusive societies, providing access to justice for all and building effective, accountable and inclusive institutions at all levels”, and for providing the evidence needed to tackle corruption-related challenges across the entire Sustainable Development Agenda.



Patrick Keuleers
Director
Governance and Peacebuilding
Bureau for Policy and Programme Support
UNDP



Hazel Feigenblatt
Managing Director, Research
Global Integrity

Introduction

UNDP's Strategic Plan, "Changing with the World" (2014-2017) places particular attention on how to better deliver results and achieve the expected impact through its programmes. To contribute to this end and in line with this, the UNDP's Global Anti-corruption Initiative (GAIN) has prioritized, as one of its main objectives, the strengthening of UNDP's anti-corruption programming. This is done by providing guidance on how to measure and assess, in a more reliable manner, the impact and results of anti-corruption interventions. This Guide responds to this objective. It presents a series of existing methodologies, tools and practices that have been used and validated by the anti-corruption community over the last few years, aiming to improve knowledge on how to make better sense of the progress in the fight against corruption.

The Guide also intends to respond to the increasing demand from a wide range of stakeholders for updated guidance on the measurement of both corruption and its absence, specifically the phenomena of transparency, accountability, and integrity. There have been several innovations in measurement approaches since the first edition of this Guide was published in 2008. More importantly, there has been a sea change in opinion as to what is most pressing in the field. We have moved from a focus on objectively and precisely measuring corruption to a focus on measuring "around" corruption for good-enough data. There has also been a collective realization that there is much less evidence on the impact of anti-corruption interventions than expected. Indeed, effectively measuring the impact of anti-corruption interventions means establishing reporting requirements and rigorous evaluation standards with quantitative, qualitative, and mixed-methods approaches.

This Guide combines a review of current measurement approaches with a discussion of challenges and concerns facing practitioners. It can be seen as a complement to the UNDP publication "**Governance Indicators: A Users' Guide,**" which is tailored to a non-specialist user and provides information on the fundamentals of indicators, data, and data collection, as well as existing data sources. This Guide is for individuals who are interested and/or engaged in anti-corruption efforts and are seeking guidance on how to think about, evaluate, and structure measurement approaches in specific contexts. Because much of the regularly collected data on corruption and anti-corruption is concerned with government activities, this Guide has a strong focus on measurement of the public sector and its relationship to civil society. It is written for practitioners involved in anti-corruption work that seek alternatives to bare logic models and input/output indicators. The Guide is divided into five chapters that cover both conceptual and practical ground.

Chapter 1 reviews the data and methods associated with corruption measurement, and existing datasets of corruption, accountability, transparency, and integrity. The second chapter introduces measurement in the monitoring and evaluation cycle, covering theories of change, measurement planning, considerations, and constraints. Chapter 3 moves on to impact and outcomes evaluation, an area that anti-corruption interventions have been failing to address. This is not necessarily because interventions are bound to fail, but rather, because measurement of the impact of complex, non-linear change processes is extraordinarily difficult with traditional means of evaluation. Chapter 4 brings practitioner voices to light, based on in-depth interviews covering the content of measurement initiatives, as well as challenges, concerns, and successes. The final chapter presents a model for good practice that involves appropriate use of methodology, stakeholder participation, and data for accountability. Throughout the Guide, mini-case studies are highlighted that illustrate the challenges or methods being discussed.

1. What is the purpose of the data that you want to collect?

	PAGE REFERENCE
Advocacy/policy reform	13, 94
Capacity building/empowerment	51, 75, 90, 112
Cross-country comparison	14, 94
Enhancing transparency, accountability, or integrity	21, 38
Impact evaluation	31, 59
Measuring corruption	11, 20, 122
Organizational performance assessment	77, 123, 124
Policy-relevance/actionability	13, 14
Project/programme evaluation	60, 63
Project/programme monitoring	31, 51
Risk assessment	70

2. What topic or field are you studying?

	PAGE REFERENCE
Accountability/oversight	25
Advocacy	69
Anti-corruption agencies	77
Capacity building/empowerment	51, 75, 90, 112
Corruption	11, 20
Institutional integrity	29, 77
Integrity	21, 29
International & regional conventions/treaties	83
National anti-corruption strategies	81
Risk	70
Rule of law	29
Sectors (education, health, forestry, etc.)	70
Social accountability	22, 27
Transparency	22

3. How can corruption and anti-corruption be measured?

		PAGE REFERENCE
Types of data	<i>Perceptions</i> <i>Experiences</i> <i>External assessment</i> <i>Administrative</i>	14
Methods of primary data collection	<i>Surveys</i> <i>Expert surveys</i> <i>Monitoring and evaluation systems</i> <i>Crowdsourcing</i> <i>Compliance/Field testing</i> <i>Indicator-driven case studies</i>	17
Using secondary data		43
Methodologies	<i>Citizen report card</i> <i>Social audit</i> <i>Community score card</i> <i>Risk assessment</i> <i>Integrity assessment</i> <i>Public expenditure tracking surveys</i> <i>Exit/User survey</i> <i>Citizen/Public official survey</i> <i>Household survey</i> <i>Expert survey</i> <i>Crowdsourcing</i> <i>Compliance/Field testing</i>	48
Design considerations	<i>Stakeholders</i> <i>De jure vs. de facto measurement</i> <i>Quality control (validity, reliability)</i> <i>Sustainability</i> <i>Constraints (costs, expertise, political will)</i>	51

4. What other aspects of measurement are you interested in?

		PAGE REFERENCE
Credibility		100
Disaggregating indicators		43
Experimental impact evaluation		59, 121
Good practices		107
Lessons Learned by practitioners		88
Proxy indicators		55, 75, 110
Results chains/Logic models		32
Theories of change		34
Theory-based evaluation	<i>Most significant change</i> <i>Outcome mapping</i> <i>Process tracing</i> <i>Contribution analysis</i> <i>Cost-benefit analysis</i> <i>Single or multiple case studies</i> <i>Systematic review</i>	63

Data, Methods, and Measurements of Corruption/Anti-corruption

This chapter provides a brief introduction to the basics of measuring levels of corruption and aspects of anti-corruption. This includes types of indicators and data, methods for data collection, and existing datasets. It also discusses the trade-offs between aggregate/composite indicators and explicit, policy-relevant indicators, and presents nuanced definitions for corruption, transparency, accountability, and integrity.

Despite the diffused nature of corruption, its impacts are felt sharply by the victims who bear its weight. In the long run, society as a whole suffers from the impact of corruption. In the short run, however, the cost of corruption in many developing countries falls disproportionately on the marginalized and vulnerable segments of society, particularly women, children, and the poor.¹

Corruption is a covert phenomenon, often unpredictable, and with characteristics that vary across time, location, and context:

- It takes many forms: *bribery, extortion, fraud, embezzlement, collusion, abuse of discretion, favoritism, gift-giving, nepotism, cronyism, patronage*
- It occurs at all levels of power: *global, national, provincial, local*
- It is conducted by agents of all types, either willingly or not: *individuals, businesses, public officials, politicians, state and non-state actors*²

Corruption is a continuously evolving phenomenon affected by various factors/determinants, which includes social and cultural settings, institutional and organizational structures, political environments, and economic and structural policies. The measurement of corruption is a means of documenting the ways in which power is subverted and resources are lost. In addition to quantifying the money lost from the public purse or bribery experiences by the public, corruption data can signal problem areas that need more attention from accountability institutions. However, corruption is hidden, and extremely difficult to capture with confidence, accuracy, or a minimal level of resources.

If corruption is defined as a misuse of government power for private gain³ – a commonly used definition of corruption – then it is a willful violation of rules or conduct. It is deliberate, illegal malpractice with the goal of personal enrichment. By contrast, the absence of corruption would involve disclosure of corruption and its risks, enforcement of rules or reform, and establishment of high standards for performance including ethics and integrity.

Measurement of anti-corruption, on the other hand, is more straightforward – it is the opposite side of corruption, and involves measuring what should exist to prevent or combat corruption. It is the measurement of transparency, accountability, and integrity within government. This includes legal and policy frameworks, institutional arrangements, processes, mechanisms, practices, outputs, and outcomes associated with these three concepts. Strong transparency, accountability and integrity mechanisms can prevent corruption, including illicit or missing flows of funds, gross abuses of power, or petty bribery by public officials.

Despite the challenges outlined above, there has been progress in measuring corruption and anti-corruption.⁴ Effective measurement builds upon well-established methods for data collection, analysis, and dissemination.⁵ One major issue that characterizes both fields of measurement is the global versus local debate, or more precisely, the measurement of national-level characteristics for comparative purposes versus the measurement of local-level contexts (province, city, community) for policy-relevant interventions.

¹Purushothaman et al., *Seeing Beyond the State: Grassroots Women's Perspectives on Corruption And Anti-Corruption*, UNDP, 2012; Carter Center. *Women and the Right of Access to Information in Liberia: A Mixed-Methods Study*. Global Access to Information Initiative, 2014.

²Phil Matsheza and Anga R. Timilsina, *Anti-Corruption Interventions for Poverty Reduction, Realization of the MDGs and Promoting Sustainable Development*, UNDP, 2008.

³Transparency International http://www.transparency.org/whoweare/organisation/faqs_on_corruption/2/. Also see Robert Klitgaard, *Controlling Corruption*, 1991.

⁴Andy McDevitt, *Gateway: Mapping the Corruption Assessment Landscape*, Transparency International, 2012.

⁵Johnson et al., *Mapping evidence gaps in anti-corruption: Assessing the state of the operationally relevant evidence on donors' actions and approaches to reducing corruption*, U4, 2012.

Table 1: Drawbacks of Aggregate Indicators

How do aggregate governance indicators fall short?
<p>Perception-based composite measures are often criticized as flawed given that perceptions are not fact and could be the reflection of distorted truth. For example, perceptions are likely to vary based on whether an expert or a common person is being asked about his/her perception about corruption. One big story of corruption in one country can easily produce a biased measure. When perception-based and experience-based surveys were compared, vast discrepancies were found between people's perceptions and people's actual experiences of corruption in a given country.</p>
<p>Comparisons over time are often problematic, as the methodologies and sources of data of many aggregate indicators change from year to year, either due to improvements in the construction of indicators or (un)availability of data sources, especially for conflict-prone or low-income countries.</p>
<p>Comparisons among countries, often done through rank ordering in aggregate indices, may also lead to false conclusions. The rank order of individual countries may take years to change, particularly if similarly-ranked countries are also undergoing reforms.⁶</p>
<p>In the same way that aggregation of many data points may fail to reflect successful reform in particular areas, it also obscures bad scores on underlying indicators.⁷</p>
<p>Another shortcoming of aggregation lies in the absence of countries from indices. Omitting a large number of countries from the rank ordering means that rankings can be misleading.⁸</p>
<p>There is the challenge of weighting the data points appropriately, which requires a clear understanding of the theoretical relationship between the topic being studied and its measurement.⁹ The assignment of additional importance to certain data points should not skew the data in a direction that is not consistent with practical realities.</p>
<p>The main criticism of aggregate indicators is that they are not easily disaggregated, let alone by income and sex, and so cannot be used to guide planning and monitoring that is sensitive to impact on women and the poor.</p>

Source: Adapted from Trapnell (2011), (World Bank 2010) and UNDP (2009a).

⁶Fredrik Galtung, "Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices," in *Measuring Corruption*, 2006.

⁷Christiane Arndt, and Charles Oman, *Uses and Abuses of Governance Indicators*. OECD, 2006.

⁸Galtung 2006; Arndt and Oman 2006

⁹Melissa Thomas, "What Do the Worldwide Governance Indicators Measure?" *European Journal of Development Research*, 2010.

Macro vs. Micro: Composite Indicators and Policy-relevant Data

Composite indicators or indices have dominated the field of corruption (and anti-corruption) measurement for a decade or more. This has happened for a variety of reasons. Corruption is a complex phenomenon to study – it is often dispersed throughout a government, from political leaders to front-line public officials, in some sectors but not others, so systemic that it seems to be everywhere, but necessarily covert so that it can't be proven anywhere. In cases like this, it is difficult to know what to measure or where to focus measurement resources. In addition, there is often a significant lag between the start of a project to the release of data, up to a year or longer. Country coverage may be limited because resources are stretched. Composite indicators gather many different data points with the goal of broad topical coverage, global country coverage, or both. For example, Global Integrity combines many data points on different areas of anti-corruption, from one collection effort, for a limited number of countries. By contrast, the Worldwide Governance Indicators combine data across many different projects to achieve global coverage. The greater number of data points can potentially result in greater accuracy, provided that the same concepts are measured consistently over time.

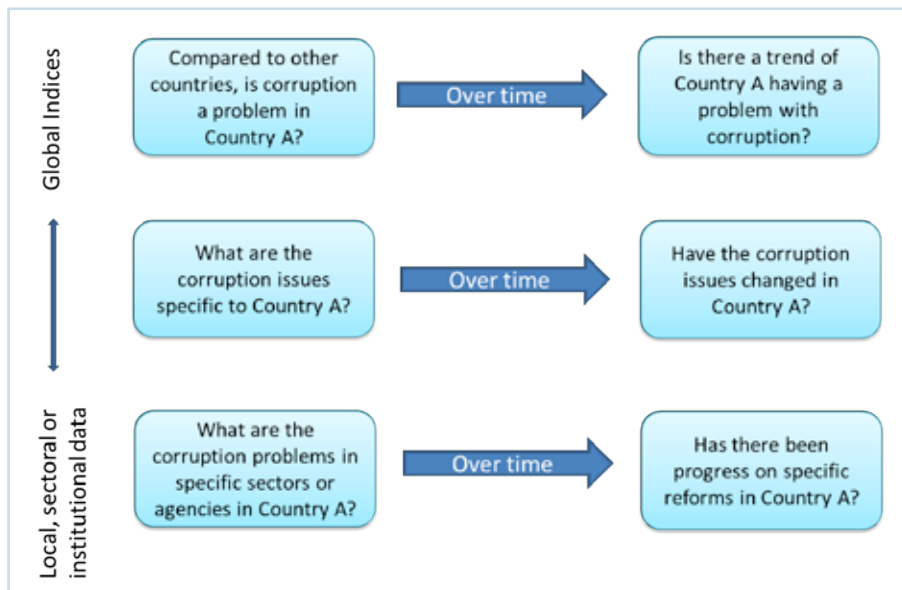
Aggregation methods pioneered in the early days of composite indicators offered a thorough and extensive means of summarizing, combining, and organizing data, and they include complementary measures of governance in a single indicator. As mentioned earlier, this was an attempt to measure complex, diffuse phenomena like corruption or governance. But there are a variety of statistical matters associated with the aggregation of data that require paying close attention before considering whether any composite indicator could be used for tracking the progress of governance or anti-corruption work. Although it is appropriate to compare high-performers and low-performers whose results fall outside of the same margin of error, there is often a missing section of middle-performers whose results are not comparable in any meaningful sense. For example, the change in the ranking or score of the Corruption Perception Index (CPI) can be as a result of one or more of the following factors: Variations in the total number of countries in the list, variations in the total number of data sources (and which ones exactly) used to calculate a country's score over the years, and change in the CPI methodology used in each year (UNDP, 2009a).

Moreover, there is also the challenge of identifying what to measure, so the resulting data will most accurately capture the extent of corruption in a given area. For example, countries that score very poorly in Transparency International's Corruptions Perception Index are seen at the top of the list compiled by the Tax Justice Network's Financial Secrecy Index.

Although various types of composite indicators may not be useful in measuring the impact of anti-corruption efforts, these measures of corruption have still been important for advocacy purposes. They allow for broad comparisons of country performance in specific sectors or areas of interest, even though comparisons must be made with care. Despite their methodological constraints, composite measures have frequently been used by governments, donors and academia. However, one of the most significant drawbacks is their broad approach to measurement. The meaning of scores or data for aggregate indicators is difficult to interpret in policy-relevant terms.

Data that has local policy relevance can be more strategically useful than composite indicators. Local data, if measured consistently over time using the same methodology, serves as the underlying basis for simple aggregation, or it can stand alone as individual data points. Datasets about local factors, sectors, and communities can demonstrate variation in outcomes within a country. This data facilitates benchmarking across provinces and within national boundaries, and provides more robust information about the local drivers of change. Yet these datasets necessarily build local context into their frameworks, thereby potentially preventing meaningful cross-country analysis. Global composite indicators provide an opportunity for benchmarking beyond country borders, but cannot provide an accurate account of cross-country comparison in terms of performance benchmarking and thus, to track the progress on any policy reforms or measure the impact of any anti-corruption intervention, practitioners require disaggregated, contextual, policy-relevant data wherever possible.

Figure 1: Global vs. Local: What Questions Can the Data Answer?



Source: Author

Figure 1 provides a snapshot of the key questions that data can answer. Global indices are more suited to answering questions about general patterns across a large number of countries, although the comparison has to be treated with caution because of the reasons mentioned earlier. As indicators become more fine-tuned, they are able to answer questions about the key issues affecting a specific country and how those issues change over time (if at all). Knowing the major problems in a country through the contextualized indicators can point to specific sectors or agencies that are underperforming, which would benefit from data pointing to reform paths on key policies or practices.

Approaches to Measurement: Data and Methods

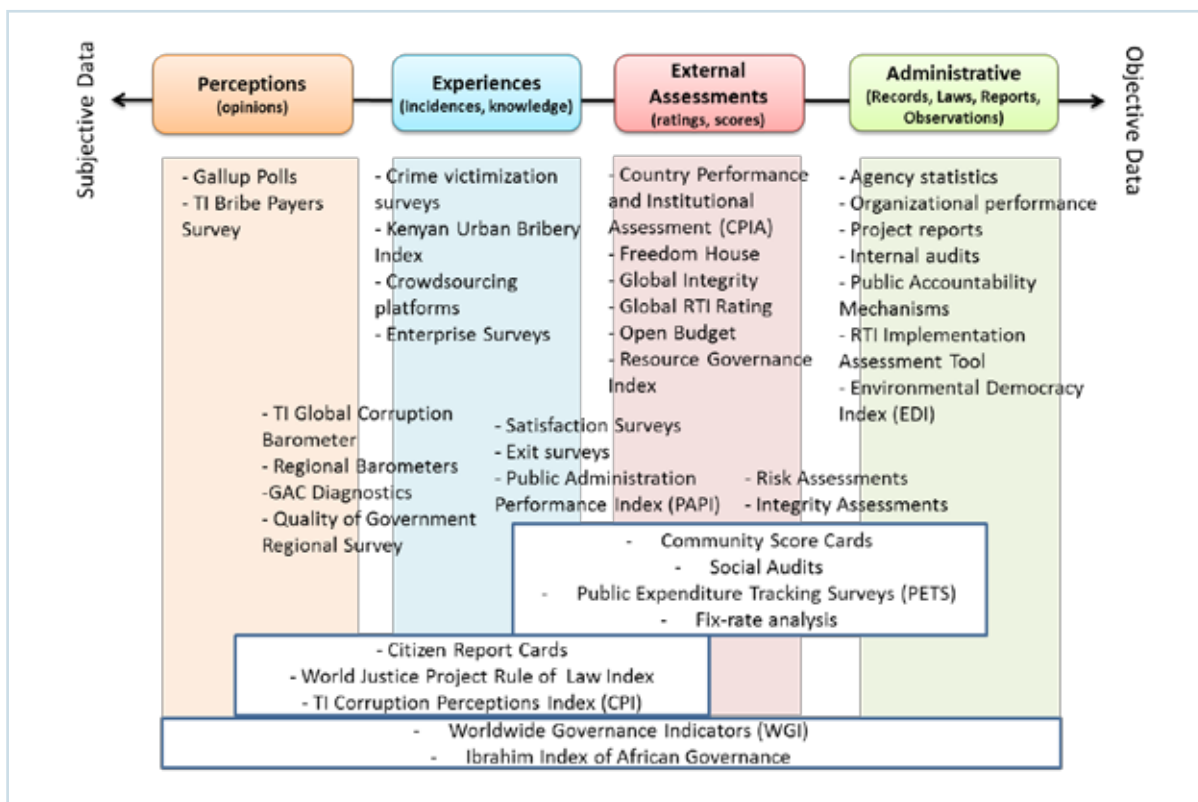
The scope and depth of measurement approaches have exploded in the last five years. There are certainly more global indices, but they cover different sectors and themes, rather than simply corruption. There are also many more innovative, country-level and community-level approaches that address both governance and corruption challenges. In fact, some of these approaches are less focused on the data than on the results of the data – what changes can be introduced based on the evidence collected? While the focus may vary from corruption to transparency, accountability, and integrity, the types of data collected and methods employed to collect these data are often common to many measurement approaches.

Types of Data

Quantitative data refers to a number, such as an amount or a score, while qualitative data is expressed in language. Often they are combined in data collection, such as when a survey asks for a rating that depends on the respondent's knowledge, opinion, or experience – this is a qualitative measure expressed quantitatively. Quantitative and qualitative are expressions of data used for different purposes. For example, for the global average of an indicator, a quantitative measure is needed. For a description of the legal framework across a region, a qualitative measure is necessary. Sometimes, they are combined in analysis: **70%** of countries have a **parliamentary** system. In any case, data can be expressed as numbers or language.

Data can also be categorized by what it represents, and this is one of the most straightforward means of understanding what the data is trying to tell you. Understanding the type of data and its limitations prevents misuse of the data, such as when perceptions of corruption are assumed to reflect actual corruption. In fact, perceptions can rise and fall independently of levels or extent of corruption. It is important to be clear about the differences in data types so that generalizations about countries are not misleading or wrong. Data on corruption and anti-corruption generally fall into four categories on a continuum from subjective (“soft”) to objective (“hard”): perceptions, experiences, assessments, and administrative data.¹⁰

Figure 2: A Continuum of Data Types Found in Selected Methods and Datasets



Source: Author

Perceptions data consists of opinions by ordinary citizens, business owners, or experts on specific topics. It is helpful for capturing information about topics that are difficult to conceptualize for objective data collection, such as public trust, civic space, grand or political corruption, and client preferences. It is also useful when administrative data is unavailable, which includes the quality of public administration or governments. Perceptions data is usually captured through surveys, and is considered a lower-cost option for collecting data, particularly if the surveys are online, through the mail, or on the telephone. But surveys of perceptions require technical competence to ensure that data is representative of the group being studied (e.g., sampling), and that the sample size is large enough to reduce margins of error. Data on perceptions is often a first step in identifying areas of focus for further exploration.

Examples: Gallup public opinion polls, Transparency International Bribe Payers Index

¹⁰These categories are particularly relevant for measurement of corruption and anti-corruption. For discussion of governance data in general, please see “Governance Indicators: A Users’ Guide,” UNDP, 2009.

Experiential data comprises specific citizen experiences (or knowledge). This includes the frequency, location, and cost of bribes, or the incidence and severity of crimes, as well as the extent of knowledge about specific laws, policies, or practices. It is useful for measuring the quality of service delivery, such as in health, education, law enforcement, and transport. It is also used to measure the extent and nature of petty corruption in particular sectors, such as bribes or crimes. It is helpful to supplement performance data collected by government agencies, which can also be used to identify bottlenecks and problems at the government-citizen interface. Experiences data is often collected through surveys, but because accuracy is important, face-to-face survey-based interviews are common. As with perceptions data, surveys of experiences may result in higher data-collection costs in order to ensure a sample size that reduces the margin of error.

Examples: Crime victimization surveys, Kenya Urban Bribery Index, Ushahidi platforms (Crowdsourcing)

External assessments are a form of data captured through scoring, rating, or ranking. The scores and ratings come from a variety of actors. Often these assessments are done by “experts” based either inside or outside the country, and these are, in fact, some of the most popular global datasets. Because of the lower costs involved with data collection and quality control (e.g., online surveys, no travel, no interviews), it is easier to cover a large number of countries. Expert assessments are often based on administrative data or third-party reports, such as case studies, audit reports, or agency statistics, and in this way can be understood as “evidence-based” assessments of corruption and governance. Assessments can also be done by ordinary citizens if they are being asked to rate or score indicators, as opposed to just giving general opinions on topics. These kinds of data-collection efforts tend to be localized to one country or even section within a country, and can provide much more specific data on petty corruption and service delivery at the country or community level. One common difference between expert assessments and citizen assessments is that the former tend to focus on country or institutional performance (e.g., grand corruption in extractive industries, transparency in public finance), whereas the latter often concentrate on micro-level impacts (e.g., petty corruption in education ministries, fraud in the provincial health system).

Examples: Freedom House, Global Integrity, Global RTI Rating, Open Budget Index, Resource Governance Index, Transparency International National Integrity System Assessment

Administrative data captures what is considered “hard measures” of government laws, activities, and performance. It often consists of agency statistics or performance data generated by governments about their own activities, as well as audit reports or project/programme reports. It is useful for assessing the quality of government resources, processes and performance. This is the easiest data to translate into action, such as reform of policy or agency practices, since data already closely adheres to existing public-sector functions. But there are questions about the reliability of self-reported data in government monitoring and evaluation systems. One complement to self-reported data by governments is data obtained through citizen feedback, observation, or in some cases through compliance or field testing by NGOs, which documents the existence, status, or completion of government activities.

Examples: Agency statistics, Organizational performance, Project reports, Internal audits, Compliance or field tests, Citizen feedback or observations

Hybrid approaches

Many methodologies combine different types of data to better capture information about the topic being studied. In reality, perceptions are often based on experience, but indicators simply capture the perception or opinion of the respondent. Assessments can be based on administrative data as evidence, but the indicators simply ask for the rating or score. The datasets produced from these indicators do not include the underlying basis for data.

By contrast, there are many datasets that comprise different types of data, as the method is very explicit. An example would be any of the Barometer data or South Korea's Integrity Assessment, as these indicators ask for *both* perceptions and experiences. Exit surveys or satisfaction surveys ask questions on what happened to the respondent, but also ask respondents to rate or score government services; the resulting data includes *both* experiences and assessments. Public expenditure tracking surveys collect experiences, assessments, and administrative data to track inputs to output stage.

Common Methods

Methods are the tools or instruments used to collect data. There are numerous methods available for the collection of data, but some are more common in the measurement of corruption and anti-corruption.¹¹ Consideration of context and research goals help determine which combination of methods is most appropriate.

Surveys can be administered to several different types of respondents, e.g., citizens, firms, and government officials, and they are helpful for capturing experiences and perceptions that point to areas of concern within public sector and governance systems. But surveys are often expensive, labor-intensive, and present technical difficulties surrounding sampling and validity. In particular, margins of error must be calculated to inform users of the accuracy of the data.

Examples: Gallup Polls, TI Bribe Payers Survey, Fix-rate analysis, Afrobarometer, Asian Barometer, Arab Barometer, Latinobarometer, AmericasBarometer, Eurobarometer, GAC Diagnostic Surveys, Quality of Government Regional Survey, National Crime Victimization Survey, Kenya Urban Bribery Index, Enterprise Surveys, Shudify, Public Administration Performance Index (PAPI)

Expert surveys, often considered a low-cost form of survey, are helpful for capturing an assessment of the nature and quality of systems. But experts are often situated in one or two locations in a country, making it difficult for data to easily capture intra-country variation. Data may also not be fully representative across the different regions of a country, unless sub-national studies are conducted.

Examples: Country Performance and Institutional Assessment (CPIA), Freedom House, Global Integrity, Global RTI Rating, Open Budget Index, Resource Governance Index, Public Accountability Mechanisms, Environmental Democracy Index (EDI), World Justice Project Rule of Law Index, TI Corruption Perceptions Index (CPI)

¹¹For additional details on the implementation of these methods, particularly surveys, see "Governance Indicators: A Users' Guide", UNDP, 2009.

Monitoring and evaluation systems embedded in government practices, as well as international development projects and programmes, may yield significant amounts of policy-relevant data, but the quality and regularity of data-collection may prevent viable cross-country comparability. It is also important to consider capacity constraints and government interest in building up data collection systems that may highlight weaknesses and inadequacies in the system. Efforts to establish common monitoring and evaluation standards facilitate the collection of comparable data.

Examples: Agency-level monitoring and evaluation systems within countries, External review and self-assessment of existing anti-corruption policies, systems and institutions, e.g., United Nations Convention Against Corruption (UNCAC), Open Government Partnership (OGP)

Crowdsourcing employs the Internet to collect and analyze information from ordinary citizens. Online crowdsourcing platforms are a rapidly growing means of collecting real-time experiences data via websites or SMS (text messages), particularly bribery incidences and election irregularities. Additionally, it may also be used to bolster transparency by aggregating citizen knowledge of policies or practices that are currently obscured, or to collect ideas from individuals outside circles of “experts” about ways to combat corruption or governance challenges.

Examples: I Paid a Bribe (example of Kallxo.com in Kosovo), Bribe Market, Ushahidi platforms, National-level reporting platforms such as “Be Responsible” in Montenegro.

Compliance review/tests are assessments performed by civil society to determine how well governments are adhering to their own rules and policies. These kinds of tests are often used to evaluate transparency and accountability systems. Right-to-information systems are often tested by information requests submitted by civil society groups, who then record details about timing delays, quality of responses, ease of appeals process, etc. Compliance tests are also employed in procurement practices to determine if information about tendering, amount of bids, and results are easily accessible by the general public.

Examples: Open Society Justice Initiative 2006, RTI Assessment & Analysis Group (RaaG) 2009; the UNCAC Compliance Review Mechanism.¹²

Indicator/Scorecard-driven case studies include interviews, document review, observation, and/or focus groups that are structured by indicators or scorecards. Qualitative methods often result in copious amounts of data that require structuring during the analysis stage, such as coding. However, indicator-driven case studies use indicators to drive data collection and lessen the need to code data after collection. The results may include narratives, qualitative data, and/or quantitative data, depending on the methodology employed.

Examples: Community Score Cards, Citizen Report Cards, Social Audits, Public Expenditure Tracking Surveys, World Bank Public Accountability Mechanisms, Carter Center RTI Implementation Assessment Tool, Public Expenditure and Financial Accountability (PEFA), Methodology for Assessing Procurement Systems (MAPS), Tax Administration Diagnostic Assessment Tool (TADAT), CRINIS (TI)

¹²The Review Mechanism of United Nations Conventions Against Corruption (UNCAC), through the self-assessment checklist filled by a state parties and a review report prepared by two reviewing countries, also provides data and information on compliance on national laws, policies and practices with the provisions of UNCAC.

Spotlight: Open Data and Corruption

Data has become much more accessible and rigorously collected in the last decade. Many practitioners are beginning to think about how to use this data for corruption measurement. The idea is that different datasets can be mashed together to provide some kind of indication on corruption.

If you are sitting in a country that has housed the majority of Transparency International surveys, it has launched an anti-corruption initiative to collect survey data on corruption, it has an “I paid a bribe” reporting mechanism, and it has an active civil society that has been collecting information on political finance contributions – what can be done with all this data?

The question becomes one of putting all of that data together and looking at it analytically. How do you make sense of it? There are different types of data, with different data structures. There’s no easy way to make it talk to each other. Together with CIVICUS, the Engine Room has been discussing ways to harmonize locally generated, citizen-reported data for comparative purposes. The study was based on a combination of desk research and mixed consultations methods with a wide group of stakeholders, including citizen representatives, national and international campaigners, data producers, government officials, donors and international experts.

Their findings are illuminating:

- There is no straightforward way to find links between locally sourced data and the large-scale corruption indices.
- Starting small and investigating specific local stories of corruption lets investigators find a thread and follow it along, slowly unraveling the complex yarn of corruption toward the bigger picture.
- Localized segmentation (where citizens look only at data directly involving them or their communities) is a boon for disentangling large lumps of data, as long as the information interests enough people to engage a groundswell of activity.

One final takeaway is that data must be used responsibly, because anyone can make data lie. It also raises questions about whether or not publishing or republishing data can actually bring about harm to individuals or to groups.

Source: “CIVICUS support to people-powered accountability and the data revolution: a scoping study by the engine room”, April 2014; “Why should we care about comparability in corruption data?” By Tin Geber, 2014: <http://schoolofdata.org/2014/05/29/why-should-we-care-about-comparability-in-corruption-data/#sthash.HNkmcj7.dpuf>

Figure 3: Tips When Assessing the Impact of Anti-corruption Work*Using indicators at the national level:*

- Nationally owned and locally produced data has more credibility than the cross-country composite indices for any policy reforms.
- Experience-based data on the impact of corruption or anti-corruption has more credibility than perception-based data.
- When there is a general lack of one type of data, the triangulation of various indicators might be useful for contextual analysis. However, tracking progress may be difficult when using a variety of indicators designed to measure different things.

Measurements of Corruption

There are many sources of data available that capture aspects of corruption, transparency, accountability, and integrity, but few datasets focus exclusively on measuring levels of corruption, except for policy experiments.¹³ It is difficult to capture precise data on corruption, hence the many datasets on *perceptions* and *experiences* of corruption.

Table 2: Selected Measurements of Corruption

Dataset	Method	Data Generated	What is being measured?
Global Integrity Report	Expert survey	Assessment	Legal framework and implementation of mechanisms for transparency, accountability, and integrity.
TI Corruption Perceptions Index	Survey	Perceptions, Experiences, Assessments	Perceptions and experiences with corruption
TI Global Corruption Barometer	Survey	Perceptions, Experiences	Perceptions and experiences with corruption
TI Bribe Payers Survey	Survey	Perceptions	Perceptions of corruption
Kenya Bribery Index	Survey	Experiences	Experiences of citizens with corruption
Enterprise Surveys	Survey	Experiences	Private-sector experiences with bribe paying to public officials in specific sectors
Governance and Corruption (GAC) Diagnostics	Survey	Perceptions, Experiences	Perceptions and experiences of corruption (administrative, state capture, bidding, theft of public resources, purchase of licenses) in the public and private sectors
Public Expenditure Tracking Surveys	Indicator-driven case studies	Experiences, Assessments, Administrative data	Flow of resources from origin to destination and determination of the location and scale of anomaly.
Afrobarometer, Latinobarometer, Asianbarometer, Arab Barometer	Citizen surveys, Household surveys	Perceptions, Experiences	Perceptions and experiences of citizens with corruption
Shudify	Exit survey	Experiences	Experiences of bribery by citizen users at local service delivery centers

¹³See Annex 3 for more details on policy experiments that use “objective” indicators of corruption.

Measurements of Anti-corruption: Transparency, Accountability, Integrity

Transparency, accountability, and integrity are often considered proxies for the *absence* of corruption, since they encompass ideas contrary to the abuse of power. These ideas include openness, responsiveness, responsibility, discipline, and ethics. But it is also useful to consider transparency and accountability as the fundamental building blocks of integrity, increasing citizen voice and participation, and influencing public officials to use power for officially authorized and publicly justified purposes.¹⁴

Transparency and accountability initiatives are a means of making government more responsive (to external actors), and responsible (through internal mechanisms), about its decision-making and activities such as the delivery of public services. Transparency can lead to accountability where officials are held responsible for their conduct and governments are held to account for their performance in both service delivery and policymaking. But accountability is not generated by the mere provision of information. It requires well-designed accountability mechanisms and integrity checks.

The 'Right to Hearing' (RTH) in Rajasthan, India has shown that effective redress of grievances is a natural and necessary next step from the Right to Information Act in pushing government from transparency to accountability. The RTH system aims to ensure a degree of accountability where the vast majority of problems with government, experienced by poor and marginalized communities, can be addressed. Rajasthan is the only State so far to have enacted a 'Right to Hearing' Act, and Rajsamand District, the only district where this Act has been implemented through an appropriate implementation system. Under the Act, citizens have the following entitlements:

1. Access to a single window system across government departments at every Panchayat for complainants to submit their complaints in writing.
2. Public hearing whereby government officials from each department are required to be present.
3. Written response within 21 days of having submitted the complaint. Automatic action should also be taken concerning officials who do *not* comply with sending a written response within the 21 days.
4. Appeals Process that allows complainants to appeal the decision outcomes of the public hearing.
5. Proactive disclosure through wall paintings of the names of beneficiaries, and a summary of benefits regarding services from every department.

One of the main impacts of the RTH system is the proactive disclosure of information for the public hearings. This creates awareness about entitlements and allows beneficiaries to monitor delivery themselves. In some cases, public hearings can serve as perfunctory social audits. Proactive disclosure is a component of the Right to Information Act and requires governments to disclose information regularly without a formal request. Proactive disclosure has also become institutionalized through public hearings even though it is not a part of the RTH Act.

Source: *Observations on the Right to Hearing Camps*,
<http://www.mkssindia.org/660/observations-on-the-right-to-hearing-camps/>

¹⁴Definition of integrity by Charles Sampford, "Understanding the Relationship between Integrity, Corruption, Transparency and Accountability," 2009.

The concepts of transparency, accountability, and integrity are broad in scope and closely interlinked, and so defy simple categorization. Transparency alone does little but inform, unless it is used to impose accountability on government for its actions. Accountability can be horizontal – oversight agencies exert influence over each other within government systems – or it can be vertical – informed citizens hold government to account through participatory processes or confrontational techniques. Integrity focuses on the systems, processes, and behaviors that exist in tension with accountability, either as a precursor, a prerequisite, or a check on unethical behavior.

The descriptions below suggest working definitions of each concept, as well as possible areas of study, methodologies, and real-life examples of available data.¹⁵

Transparency

Transparency can be defined as the availability of information, both to the general public and to individuals that comprise the government workforce, as well as clarity about government processes, rules, and decisions.

Right-to-information systems are practical components of government administration that reflect a commitment to the principle of transparency. Although right-to-information systems comprises only one part of transparency in government, they are a key factor in bringing about openness and access to information. A right-to-information system aims to increase the transparency of government by providing regular and reliable information to the public and facilitating appropriate and relevant use of that information. Monitoring the practices of a right-to-information system can identify whether problems are rooted in a lack of political will or the capacity of an administration. Measurement of right-to-information systems might entail data on legal frameworks, institutional arrangements for monitoring and oversight, records and information management, information disclosure procedures, and proactive disclosure outputs, including open-data resources.

Table 3: Selected Examples of RTI Measurements

Dataset	Method	Data generated	What is being measured?
Global Integrity Report (National and Subnational)	Expert survey	Assessment	Legal framework for right to information and administrative practices for requests and appeals, whistleblower protections
World Bank Public Accountability Mechanisms	Expert survey, Public officials survey, Indicator-driven case studies, Compliance testing	Assessment, Administrative data	Legal framework for right to information, administrative practices for requests and appeals, monitoring and oversight arrangements, proactive disclosure outputs, response and appeal rates, strength of enabling environment (civil society, media)
Centre for Law and Democracy/AccessInfo Europe Global RTI Rating	Expert survey	Assessment, Administrative data	Legal framework for right to information
Carter Center Implementation Assessment Tool	Indicator-driven case studies	Assessment	Administrative practices for requests, administrative resources and capacities

¹⁵Individual data sources may cover various aspects of transparency, accountability, and integrity within one dataset, and therefore may appear in more than one table below. Attention should be paid to the column "What is being measured" for specific information on topics covered.

Table 3: Selected Examples of RTI Measurements *(continued)*

Dataset	Method	Data generated	What is being measured?
OECD Government at a glance : Transparency in Governance	Public officials survey	Administrative data	Legal framework for right to information, proactive disclosure outputs
Open Society Justice Initiative: Transparency and Silence	Compliance testing, Indicator-driven case studies	Administrative data	Response rates
RTI Assessment & Analysis Group – RaaG (India)	Compliance testing	Administrative data	Response rates
World Wide Web Foundation Open Data Barometer	Expert survey, Compliance testing	Assessment, Administrative data	Proactive disclosure (open data) outputs, administrative resources and capacities, impacts of open data
Open Democracy Advice Center Golden Key Awards (South Africa)	Compliance testing, Indicator-driven case studies	Assessment, Administrative data	Response and appeal rates, Administrative practices for requests, administrative resources and capacities
Open Knowledge Foundation Open Data Index	Expert survey, Compliance testing	Assessment, Administrative data	Proactive disclosure (open data) outputs
Transparency International: Alternative to Silence	Expert survey	Administrative data	Whistleblower protections
Country-level Monitoring and Oversight	Monitoring & evaluation systems	Administrative data	Response and appeal rates

The Open Democracy Advice Center The Golden Key Awards (GKA) recognize best practices in the implementation of the South African RTI law (PAIA), and acknowledge entities and organizations that model openness, responsiveness and information sharing. In the case of the public sector the following sources were consulted: Institution websites, PAIA reports to the Ombudsman, Section 14 PAIA manuals, Internal PAIA procedures, Filing Plans, and Information Officer Questionnaires. After a six-week data-collection period, the research team analyses the data and together with the panel of judges, scores each institution. The research aims to ascertain the internal readiness of institutions to implement PAIA. Areas of evaluation during the research include policy, compliance, records management and the availability of human and capital resources for the administration of PAIA to determine the readiness of institutions.

The awards are given under the following categories:

- The openness and responsiveness award by institution;
- Deputy Information Officer of the year award;
- Requester award; and
- The best media usage /engagement with PAIA.

The objective of the awards goes beyond recognizing practice; it aims to encourage non-performing institutions to accelerate their performance and possibly stand a chance of nomination and winning an award.

Source: South African Human Rights Commission, <http://www.sahrc.org.za>

Public-sector functioning in procurement, tax, revenue, budgets, and human resources are important areas for transparency in central government activities. Disclosure or information dissemination is often part of public-sector reform efforts that aim for more effective resource allocation or gains in efficiency and productivity. Better resource allocation in budgets is often the primary goal of fiscal transparency efforts, including policies that are pro-poor, gender-sensitive, and responsive to marginalized groups. But efficiency gains and cost savings can also be realized by institutionalizing transparency in budgeting processes and procedures. Cost- and time-saving results in procurement practices can be obtained through openness during the bidding process and technological streamlining of activities.

Transparency in procurement and revenues also facilitates monitoring of activities, which helps to ensure openness and transparency throughout the procurement process, including tendering, qualifications, and awards. Transparency in human resources management is related to information disclosure of policies, procedures, and practices of personnel. A lack of transparency in the wage bill or in personnel rosters may obscure the presence of “ghost employees” that represent a substantial percentage of salary costs. Measurement of transparency in public-sector functioning would focus on elements of openness and access in procurement, tax, revenue, budgets, and human resources.

Table 4: Selected Examples of Public-Sector Transparency Measurements¹⁶

Dataset	Method	Data generated	What is being measured?
International Budget Partnership Open Budget Index	Expert survey	Assessments, Administrative data	State of budget transparency, participation, and oversight
IMF Reports on the Observance of Standards and Codes	Indicator-driven case studies	Assessments, Administrative data	Extent to which countries observe certain internationally recognized standards and codes, including data dissemination; fiscal transparency; and monetary and financial policy transparency
Public Expenditure Tracking Surveys	Indicator-driven case studies	Experiences, Assessments, Administrative data	Flow of resources from origin to destination and determination of the location and scale of anomaly.
Revenue Watch Institute Resource Governance Index	Expert survey	Assessments	Quality of governance in the oil, gas and mining sectors, including institutional & legal setting and reporting practices (transparency)
Public Expenditure and Financial Accountability (PEFA)	Indicator-driven case studies	Assessments, Administrative data	Capacities and performance of national public financial management systems, including fiscal, tax and procurement transparency
OECD Methodology for Assessing Procurement Systems (MAPS)	Indicator-driven case studies	Assessments, Administrative data	Capacities and performance of national public procurement system, including transparency
Tax Administration Diagnostic Assessment Tool (TADAT)	Indicator-driven case studies	Assessments, Administrative data	Capacities and performance of a country's system of tax administration, including transparency

¹⁶See Annex 4 for additional transparency indicators at the national and organizational level.

International Budget Partnership: Open Budget Index

The Open Budget Survey is a comprehensive and analytical survey that evaluates whether governments give the public sufficient access to budget information and opportunities to participate in the budget process at the national level. The IBP works with civil society partners in 100 countries to collect the data for the Survey. The first Open Budget Survey was released in 2006 and is conducted biennially.

To easily measure the overall commitment of the countries surveyed to transparency and to allow for comparisons among countries, IBP created the Open Budget Index (OBI) from the Survey. The OBI assigns a score to each country based on the information it makes available to the public throughout the budget process.

In September 2014, IBP launched the Open Budget Survey Tracker, an online monitoring tool allowing citizens, civil society, media, and others to monitor in real time whether central governments are releasing the requisite information on how the government is managing public finances.

The Open Budget Index correlates significantly with the human development index, gender-related development index, gender empowerment measure, primary education, and water and sanitation access. However, when differences in per-capita income and region are held constant, budget transparency retains a significant statistical association with only a few variables, namely infant and child survival, the percentage of the population using improved drinking water, and public health expenditure levels.¹⁷

Source: International Budget Partnership, <http://internationalbudget.org/>

Accountability

Accountability comprises “answerability,” which refers to the rights of citizens to request a response to questions about government decision-making, as well as the obligation of government to respond. It also includes “enforcement,” which is about the capacity to ensure that action is taken, and provides access to mechanisms for redress when accountability measures fail.¹⁸ For governments, accountability is manifest in its role in ensuring effective delivery of services (e.g., education, health, social welfare, transportation, etc.) and as a protector of the public interest (e.g., law, order, security, safety, etc.). Accountability may also be instituted as an organization holding itself to account through internal audits, oversight mechanisms, and risk assessments.

In terms of accountability in the public sector, the following four pillars have been often cited:

- financial accountability: reports on the intended and actual use of resources or of designated offices
- administrative accountability: critical systems of control internal to the government
- political accountability: free and transparent elections as an effective starting point for oversight
- social accountability: civic engagement that involves ordinary citizens and groups demanding greater accountability for public actions and outcomes¹⁹

¹⁷Fukuda-Parr et al., *Does Budget Transparency Lead to Stronger Human Development Outcomes and Commitments to Economic and Social Rights?* International Budget Partnerships International Budget Partnership Working Paper, 2011.

¹⁸Andreas Schedler, “Conceptualizing Accountability.” In *The Self-restraining State: Power and Accountability in New Democracies*, 1999.

¹⁹UNDP, “Mainstreaming Anti-Corruption in Development”, Anti-corruption Practice Note, 2008.

In terms of a broad categorization, the literature often discusses horizontal and vertical accountability. **Horizontal accountability** occurs through oversight by government agencies such as Parliament, Ombudsman, Supreme Audit Institution, and the Anti-corruption Agency. These agencies are authorized to conduct audits, investigations, and enforce sanctions for failure to meet standards. Models of this type of accountability focus on a theory of change that involves setting standards (*see Integrity*), obtaining information about behavior (*see Transparency*), making judgments about whether behaviors violated accepted norms, and applying effective sanctions for violations.²⁰ Measurements of horizontal accountability might include institutional performance indicators of oversight agencies (*See Section 3*) and tracking violations and sanctions. **Institutional accountability** is implemented through mechanisms such as internal and external audits, appeals, oversight, and legislative scrutiny of public sector functions. Risk assessments may also be included as a measurement of accountability.

Table 5: Selected Examples of Measurements of Institutional or Horizontal Accountability²¹

Dataset	Method	Data generated	What is being measured?
International Budget Partnership Open Budget Index	Expert survey	Assessments, Administrative data	State of budget oversight
IMF Reports on the Observance of Standards and Codes	Indicator-driven case studies	Assessments, Administrative data	Extent to which countries observe certain internationally recognized standards and codes, including auditing; anti-money laundering and countering the financing of terrorism (AML/CFT); banking supervision; corporate governance; insurance supervision; payments systems; and securities regulation
Revenue Watch Institute Resource Governance Index	Expert survey	Assessments	Quality of governance in the oil, gas and mining sectors: including Institutional & Legal Setting and Safeguards and Quality Controls
Public Expenditure and Financial Accountability (PEFA)	Indicator-driven case studies	Assessments, Administrative data	Performance of national public financial management systems, including fiscal, tax and procurement audits, oversight, and accountability practices
OECD Methodology for Assessing Procurement Systems (MAPS)	Indicator-driven case studies	Assessments, Administrative data	Performance of national public procurement system, including audits, oversight, and accountability practices
Tax Administration Diagnostic Assessment Tool (TADAT)	Indicator-driven case studies	Assessments, Administrative data	Performance of a country's system of tax administration, including audits and accountability practices
World Bank Public Accountability Mechanisms	Expert survey	Administrative data	Strength of legal framework for oversight of conflict of interest safeguards, financial disclosure, and immunity protections
External audits (usually performed by Supreme Audit Institution)	Indicator-driven case studies, Monitoring & evaluation systems	Assessments, Administrative data	Report on an organization's accounts and financial statements, the legality and regularity of its operations, and its financial management procedures and financial performance
Internal audits (performed by agency on itself)	Indicator-driven case studies, Monitoring & evaluation systems	Assessments, Administrative data	Self-assessment on various aspects of an organization's accounts and financial statements, the regularity of its operations, and its financial management procedures and financial performance

²⁰Mick Moore and Graham Teskey. *The CAR Framework: Capability, Accountability, Responsiveness. What Do These Terms Mean, Collectively and Individually?* A Discussion Note for DFID and Conflict Advisers, 2006.

²¹See Annex 5 for additional accountability indicators at the national and organizational level.

Vertical accountability encourages civil society to demand that governments rectify problems efficiently and effectively, even though they have little force of sanctions. Vertical forms of accountability are advanced by civil society organizations through various means of protest and advocacy. Meanwhile, diagonal accountability arrangements see citizens engage directly with horizontal accountability institutions through policymaking, expenditure tracking, and participatory budgeting processes.²²

These mechanisms can be participatory or confrontational, depending on the receptiveness of the government to inspection and reform, and the ability of civil society to bring the problem to light and seek redress. There is an element of change since the government is expected to rectify the problem. Service delivery in education, health, forestry, water, environment and other sectors has been a central focus of transparency and accountability initiatives, primarily because of the poor record of government action and the presence of corruption in weak institutional environments. Sharing of information between government and beneficiaries concerning performance and capacities, as well as the supply of on-the-ground information from beneficiaries, can lead to better prioritization of goals. It can also build cooperative partnerships and improve effectiveness in the delivery of services. Transparency initiatives in service delivery are also important for accountability purposes, and frequently appear in demand-side efforts that target increased effectiveness, or efforts to curb corruption in public management systems. Improved decision-making in all areas of the public sector can be facilitated by increased openness across departments/units/agencies, increased access to information for the public, and increased receptiveness to externally-generated information. Measurement of vertical accountability includes monitoring of government activities, evaluation of government outputs, and tracking of outcomes in the community.

Grassroots Women Fighting Corruption

When grassroots women organize at the community level to design and implement anti-corruption strategies, they build gender sensitive governance that leads to improved service delivery, increased access to justice, and decreasing levels of corruption and poverty.

Through the Huairou Commission and UNDP PACDE's Transparency & Accountability Initiative, six grassroots women's organizations designed and implemented anti-corruption projects in the sectors of health care, water and sanitation, electricity, land, and national identification documents. 179 grassroots women were mobilized to lead these projects, impacting 2,338 community members across Uganda, the Philippines, Nicaragua, Brazil, and Nepal, while 508 community members were trained in technical skills.

These pilots, highlighting the link between organizing and partnerships at the local level, and governance and service delivery, have led to an increased understanding of the grassroots women's approach to anti-corruption. Such an approach is collective, goes beyond confrontation, and is rooted in improving governance and transforming traditional power dynamics.

Source: Huairou Commission, <http://huairou.org/transparency>

²²Jonathan Fox, "Social Accountability: What Does the Evidence Really Say?" World Bank, 2014.

Table 6: Selected Examples of Vertical Accountability Measurements

Dataset	Method	Data generated	What is being measured?
Global Integrity Report (Papua New Guinea Provincial Healthcare)	Expert survey	Assessment	Legal framework and administrative practices in healthcare transparency and accountability
Shudify	Exit surveys	Assessment, Experiences	Public-sector performance, user experiences with bribery
Punjab Model for Proactive Governance	Exit surveys	Assessment, Experiences	Public-sector performance, user experiences with bribery
Public Expenditure Tracking Surveys	Indicator-driven case studies	Assessment, Administrative data	Allocation of resources and potential weaknesses in the mechanisms used to allocate resources
Community-level Community Scorecards, Citizen Report Cards, Social Audits	User surveys, Public officials surveys, Indicator-driven case studies	Assessment, experiences, Administrative data	Effectiveness of basic services such as education, health, and water and sanitation
Fix-rate	User surveys, Crowdsourcing, Indicator-driven case studies	Perceptions, Experiences	Incidence with which transparency and accountability problems are resolved to the satisfaction of key stakeholders
Country-driven monitoring and oversight	Monitoring & evaluation systems	Administrative data	Effectiveness of basic services such as education, health, and water and sanitation

Tamasha: Young People Checking if Governments and Services Work for Them

“Though young people aged 10-29 years constitute a large portion of the population in East Africa, their perspective is often absent in mainstream discourse. In 2010 the African NGO Twaweza supported the youth focused organization Tamasha to undertake in-depth monitoring in 32 communities in 8 districts in Tanzania.

A total of 960 households were surveyed, educated and empowered to think and act on issues affecting their rural communities, and 595 youth facilitators were recruited, trained and engaged in each respective district to implement the programme.

In each community one male-one female pair of youth were involved in monitoring service delivery and the treatment of youth over a minimum period of one week, including through the use of ‘dummy patient’ technique. Quantitative and qualitative information (‘stories’) were collected, and both audio-recordings and photographs were captured. The fieldwork was completed in 2010, with the results analyzed and the main report, popular briefs and exhibition launched in 2011.” Outputs of the study include: a research report, policy briefs, and presentation on health, utilities, livelihoods, youth and services in English and Swahili, the dissemination of research through Facebook, public forums, radio, television, and print media, and a depository of compiled research analysis and photographs.

The Rule of Law can be defined as the restriction of the arbitrary exercise of power by subordinating it to well-defined and established laws.²³ In practice, the rule of law is a system of rules and rights that enables fair and functioning societies.²⁴ It is the ultimate form of accountability, whereby individuals are detained and investigated by law enforcement, held to account before a court of law, and punished accordingly. Measurement of the rule of law might focus on the strengths and effectiveness of law enforcement, and judicial and correctional institutions. It might also include notions of justice, openness, and access to judicial services for poor and marginalized groups.

Table 7: Selected Examples of Rule of Law Measurements

Dataset	Method	Data generated	What is being measured?
World Justice Project Rule of Law Index	Expert survey, Household survey	Perceptions, Experiences, Assessments	Extent to which countries adhere to the rule of law on several dimensions: constraints on government powers; absence of corruption; order and security; fundamental rights; open government; effective regulatory enforcement; access to civil justice; effective criminal justice; and informal justice
CLEEN Foundation (Nigeria)	Indicator-driven case studies, Citizen surveys	Perceptions, Experiences, Assessments, Administrative data	Police violence, crime victimization, police discipline, police accountability, etc.
European Commission for the Efficiency of Justice (CEPEJ)	Public officials survey (Member states)	Administrative data	Evaluation of Judicial Systems, Judicial time management, Quality of justice, Enforcement, Mediation
UNODC Statistics	Public officials survey (Member states)	Administrative data	Statistics on drug use and trafficking, criminal justice, and crimes
UNCAC Review Mechanism Reports	Indicator-driven case studies	Assessments, Administrative data	Extent to which signatory parties comply with the provisions of the convention on legal and regulatory regimes to fight corruption
Country-driven court, crime and police statistics	Monitoring & evaluation systems	Administrative data	Various data on crimes, law enforcement activities, court processes, investigations, prosecutions, convictions, etc.

Integrity

Integrity as it is used in the fields of corruption and anti-corruption has taken on a two-tiered meaning. At the national level, integrity refers to the strength and effectiveness of several pillars of a country's governance system. Together, these pillars manage corruption risks and contribute to the fight against corruption: Legislature, Executive branch of government, Judiciary, Public sector, Law enforcement, Electoral management body, Ombudsman, Audit institution, Anti-corruption agencies, Political parties, Media, Civil society, and Business.²⁵ Along with the national-level integrity, practitioners often use the concept of institutional integrity, such as judicial integrity, referring to the integrity of a particular institution or a sector.

²³Oxford Dictionary.

²⁴World Justice Project, <http://worldjusticeproject.org/what-rule-law>

²⁵National Integrity Systems (NIS) Assessment, Transparency International, <http://www.transparency.org/whatwedo/nis>

At the individual level, integrity refers to ethical behavior and personal responsibility that all public officials must embody. These models are reflected in conflict of interest restrictions that allow government organizations to hold their employees to a certain standard expected of public office holders. Similarly, financial disclosure systems monitor the behavior of officials for conflicts that may compromise their integrity or instances of illicit gain. Measurement of integrity can extend from the rules that establish codes of conduct to the violations of those rules through petty corruption.

Table 8: Selected Examples of Integrity Measurements

Dataset	Method	Data generated	What is being measured?
Global Integrity Report (National and Subnational)	Expert survey	Assessment	Legal framework and administrative practices related to conflicts of interest, financial disclosure, and accountability practices for legislature, executive branch, and judiciary
World Bank Public Accountability Mechanisms	Expert survey	Assessment, Administrative data	Legal framework for conflict-of-interest safeguards, financial disclosure, and immunity protections; monitoring and oversight arrangements, sanctions
OECD Government at a glance : Transparency in Governance	Public officials survey	Administrative data	Legal framework for conflict of interest safeguards and financial disclosure
Country-level reports on judicial integrity	Varies	Varies	Various dimensions of judicial integrity, impartiality, access to justice, and effectiveness of courts
Agency-level Integrity Assessments	Varies	Varies	Various dimensions of vulnerabilities to corruption, actual levels of corruption, organizational ethics and administrative culture
Transparency International National Integrity Systems Assessment	Indicator-driven case studies	Assessments, Administrative data	Legislature, Executive branch of government, Judiciary, Public sector, Law enforcement, Electoral management body, Ombudsman, Audit institution, Anti-corruption agencies, Political parties, Media, Civil society, Business
Korea Anti-Corruption & Civil Rights Commission (ACRC) The Anti-Corruption Initiative Assessment (AIA) at the institutional level	Expert opinions and scores	Assessment of institutional measures/initiatives	Anti-corruption will & efforts (e.g., establishment of anti-corruption systems, enhancement of policy transparency & reliability, reduction of corruption risks, promotion of a culture of integrity in the public sector) and anti-corruption achievements

Measurement and the M&E Cycle

This chapter addresses the role of measurement in the cycle of monitoring and evaluation, particularly in building theories of change, designing indicators, and finding or collecting data. It also distinguishes between methods for data collection, and methodologies, which consist of the activities that drive a measurement initiative toward its goals. The chart of methodologies provides information on the nature and purpose of various measurement approaches that can be used in the M&E cycle. There is also a discussion of the various constraints that must be considered with planning measurement initiatives.

Measurement of anti-corruption interventions may involve the measurement of levels of corruption or aspects of anti-corruption, as discussed in the previous chapter. But the measurement of projects, programs, or activities is primarily concerned with three key processes: assessing whether objectives are being met, evaluating the quality of performance, and estimating the extent of influence on external outcomes. Surrounding these processes are questions about timing, causes, agents, locations, and mechanisms of change. Monitoring and evaluation approaches attempt to answer those questions in rigorous and practical ways so that data can be used effectively for the good of the project. Moreover, monitoring and evaluation could vary in scope depending on whether one is evaluating a programme/project, the capacity and performance of an institution, the effectiveness of the overall national integrity system, or the outcomes of national anti-corruption strategies or policies.

Monitoring and evaluation (M&E) are often conducted at different points during the project cycle, but in recent years, they have been combined in new approaches that address more complex, actor-oriented interventions.

Monitoring is an ongoing process of obtaining feedback on how well a project or activity is on track to meet its objectives.

Monitoring processes can focus on compliance, activities, finances, organizational capacities, beneficiaries and more commonly, results.

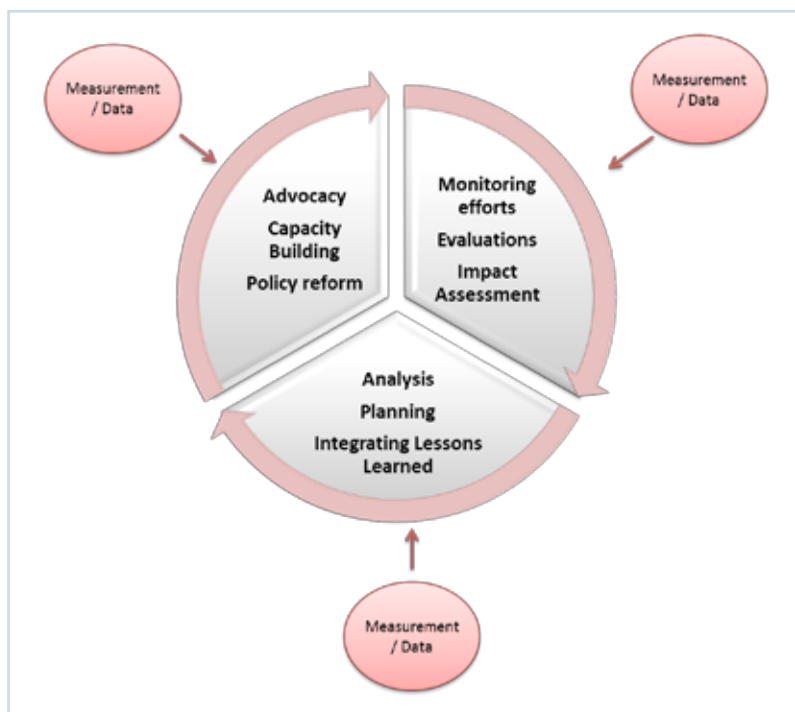
Evaluation is a means of establishing the efficiency, sustainability, value, or relevance of a project.

It is often conducted mid-cycle or end-of cycle as a comprehensive assessment of project performance thus far, whereas monitoring is ongoing throughout the life cycle of a project.

Impact evaluation (or impact assessment) is a type of evaluation that involves understanding the nature of the change that has taken place, including any negative or unintended consequences.

It is an assessment of completed activities in order to attribute causality or determine the extent of contribution to external outcomes i.e., the effectiveness of an activity or project at reducing corruption or enhancing transparency, accountability, and integrity.

Figure 4: Relationship Between Measurement/Data and M&E Processes



Source: Author

Measurement and data feed into all stages of the M&E cycle, from planning activities to monitoring and evaluation. In fact, it is important to design monitoring and evaluation approaches during the project planning process because of the close relationship between project and M&E goals. The purpose of M&E is to track progress on project objectives and assess the contribution of project results on external outcomes and other quality control elements. The planning process thus includes conducting analysis, developing the results logic, defining clear results and indicators, measuring baselines and setting targets.²⁶

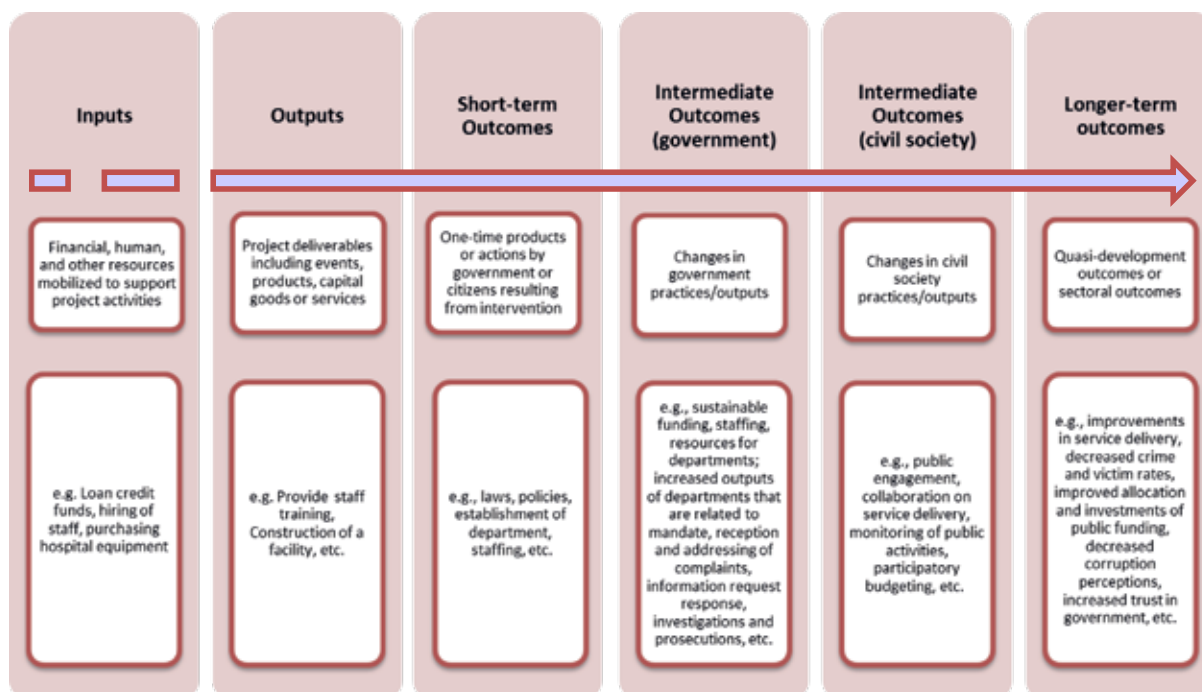
Logic Models

Logic models illustrate program components and help stakeholders clearly identify inputs, outputs, and outcomes. They focus on the expected causal links—the “program logic”—in the following results chain: inputs, outputs, outcomes, and impacts (such as long-term outcomes).²⁷ They emphasize the development of indicators, particularly at the outcome stage, as this is where changes in behavior are most expected.

²⁶UNDP *Handbook On Planning, Monitoring And Evaluating For Development Results*, 2009.

²⁷World Bank, *Monitoring and Evaluation: Some Tools, Methods, and Approaches*, 2004

Figure 5: Results Chain for Corruption and Anti-corruption Interventions



Source: Author

In terms of indicators for results chains, relying on levels of corruption to indicate change or impact is not optimal. There are multiple factors that feed into the levels, nature, amount, or extent of corruption. It is rarely possible to establish links of causality from reform action to broad corruption measurement. Instead, tracking progress should be conducted at lower levels of analysis, where it is possible to understand why change is happening.

“Gauging and measuring impact is not just about measuring corruption but also measuring the confounding factors. So we can measure corruption, but what drives it? It’s a research problem. Measuring impact is many steps removed from measuring corruption.”²⁸

The solution is to focus on approaches that allow changes to be tracked. Once data is collected, an investigation should be conducted on whether the reform efforts are linked to measurable progress.

²⁸Interview with Mihaly Fazekas, University of Cambridge, 2014.

Theories of Change

Logic models or frameworks focus explicitly on the results chain from inputs to outcomes (and impacts). They fail to take into consideration external factors that may either influence outcomes or are beyond the control of project parameters.

***Theories of Change** add elements of explanation and consideration of external influences to the results chain. They link outcomes and activities to explain how and why the desired change is expected to happen. Theories of change also make explicit the political and economic processes that serve as both preconditions to project activities as well as confounding factors in the process of change.*

The historically dominant principal-agent model of ‘grand theories of change’²⁹ sees relationships in terms of oversight, incentives, and expectations: Does an agent (government) act in the best interest of a principal (citizen) who has conferred upon him some decision-making authority? In contrast, the collective-action model views accountability as a problem of non-action among a group of individuals with common interests. In order to counter systemic corruption, a certain threshold of action must happen to tip the system into accountability. Focusing on one aspect of the problem will not suffice.

“It’s a network of actors, organizations, persons, so just taking that one out from the network doesn’t change the whole structure. It doesn’t fall apart. If you want to change the logic of the system, intervening here and there doesn’t work. You have to change the whole network.”³⁰

These models, in fact, are complementary depending on the context. Mobilizing citizens and revamping entire power structures are not replacements for establishing clear lines of vertical or horizontal accountability within government. A recent study of social accountability projects revealed that governance reforms that coordinate citizen voice with government capacity are more promising. Strategic approaches to accountability will coordinate citizen voice initiatives with governmental reforms that bolster public-sector responsiveness.³¹

Countering corruption is about changing behaviors, attitudes, and the structures of accountability, transparency, and integrity. This kind of change is a contentious, political process. Linear models of change that rely on a stable set of outputs are rarely adequate to capture the multiple actors, relationships, and behaviors that characterize anti-corruption efforts. Jonathan Fox argues that “when accountability efforts actually work, it is often because initiatives in one arena trigger pro-accountability actions in another (as when electoral pressures or citizen action kicks checks and balances into gear).”³² The implications are profound: Anti-corruption efforts are not bound by location, time, or actor. Moreover, corruption has multiple causes (direct and indirect) and manifests differently according to context. Consideration of these complexities is required to identify a causal process and consistent theory of change.

Theory of change is a way to capture the role of the collective action model. But many anti-corruption interventions do not have a built-in theory to guide the results chain, making it difficult to meaningfully track the progress and measure the desired results. Theory of change can serve as a conceptual map of the change process from start to finish. It makes known the underlying assumptions about why and how a project will be successful, and maps out the intermediate steps that must be taken to reach a long-term result. An explicit theory of change outlines a robust framework for monitoring, and more importantly, leaves an evidence trail for evaluation purposes.

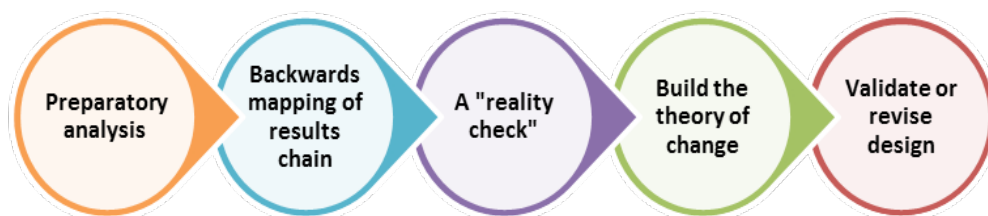
²⁹Jesper Johnson, *Theories of change in anti-corruption work: A tool for programme design and evaluation*, Chr. Michelsen Institute (CMI), 2012.

³⁰Interview with Mihaly Fazekas, University of Cambridge, 2014.

³¹Jonathan Fox, *Social Accountability: What does the evidence really say?*, World Bank, 2014.

³²Fox 2014.

Figure 6: Five-step Approach to Building a Theory of Change in Anti-corruption Interventions



Source: Jesper Johnson, *Theories of change in anti-corruption work: A tool for programme design and evaluation*, Chr. Michelsen Institute (CMI), 2012.

The theory of change process for anti-corruption interventions starts with an analysis of the political and economic processes that prevail in the project context, including the incentives, relationships, distribution and contestation of power among stakeholders. The process then moves to identifying long-term goals, and mapping out the steps needed to reach those goals. The “reality check” is in considering the internal logic of the results chain, and the external factors that may influence outcomes. Building the theory of change requires identifying where preconditions are necessary for plausible causal pathways to occur, as well as the challenges that may prevent goals from being achieved (shown as red boxes in Figure 5). Based on the outcomes identified, and given the prevailing political and economic constraints, indicators and targets are developed to monitor performance and assist with evaluations.

Figure 7: Theory of Change Analysis for Anti-corruption Enforcement Work

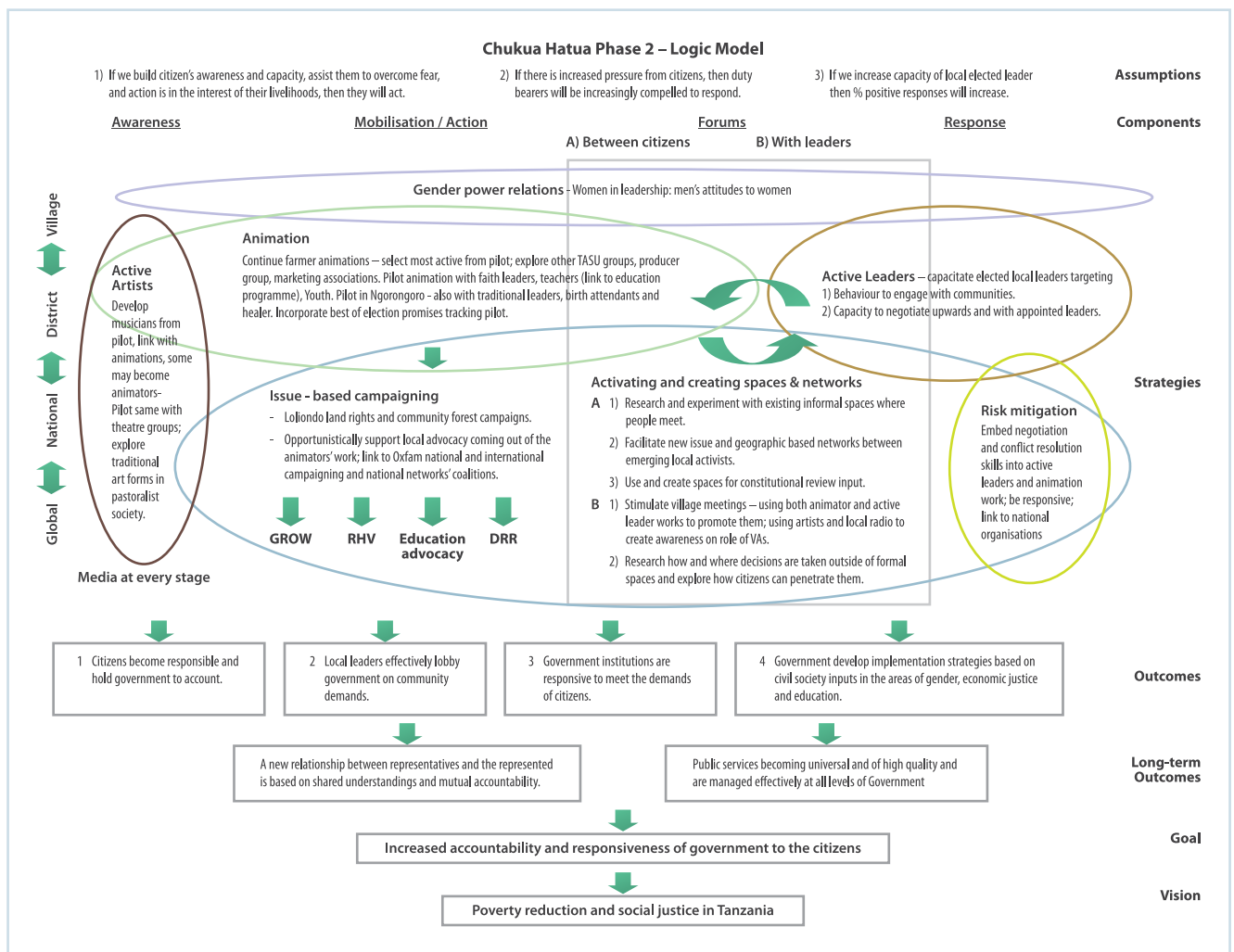


Source: Adapted from Jesper Johnson, *Theories of change in anti-corruption work: A tool for programme design and evaluation*, Chr. Michelsen Institute (CMI), 2012.

Logic models are often incorporated into theories of change, allowing the development of indicators to track progress. The indicators developed through a theory of change process should meet certain SMART conditions in order to serve as effective measurements of results:

- Specific:** Is the indicator specific enough to measure progress toward the results?
- Measurable:** Is the indicator a reliable and clear measure of results?
- Attainable:** Are the results in which the indicator seeks to chart progress realistic?
- Relevant:** Is the indicator relevant to the intended outputs and outcomes?
- Time-bound:** Is data available at reasonable cost and effort?

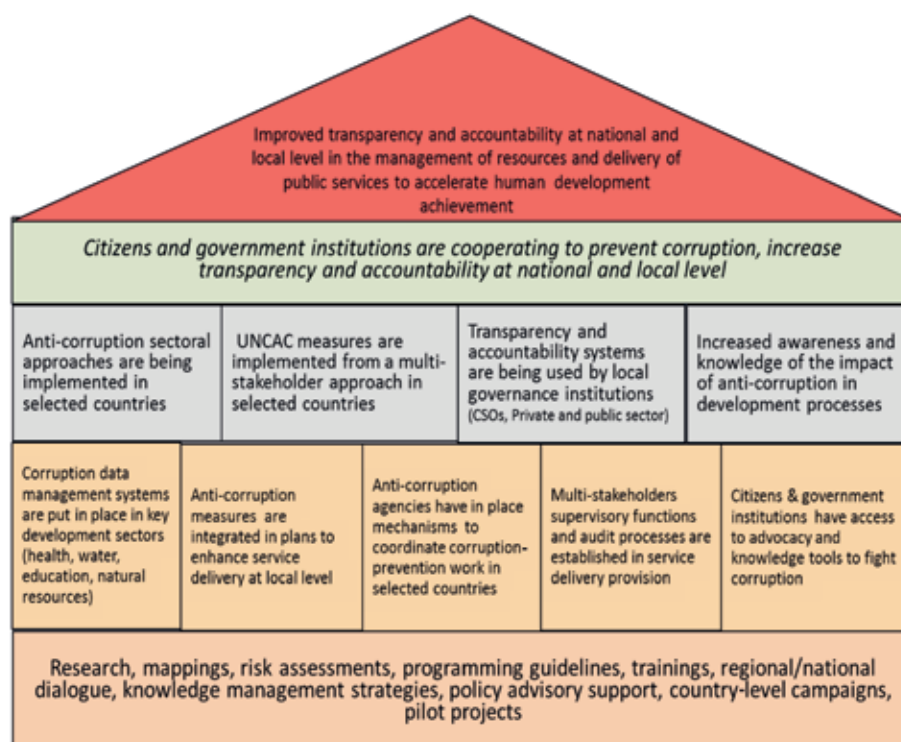
Figure 8a: Logic Model with Theory of Change



Source: Richard D. Smith and Dunstan Kishekya, Chukua Hatua, Tanzania Effectiveness Review, Oxfam, 2013.

An illustrative example is presented below to show that overall engagement on anti-corruption should aim to contribute to improving transparency and accountability by using a theory of change.

Figure 8b: Logic Model with the Theory of Change (An Example from UNDP's Global Anti-corruption Initiative)³³



The process of formulating indicators should begin with the following questions:

- How can we measure and track the expected results being achieved?
- What type of information demonstrates a positive change?
- What can be feasibly monitored with given resources and capacity constraints?
- Will timely information be available for the different monitoring exercises?
- What will the system of data collection be and who will be responsible?
- Can national systems be used or augmented?
- Can government indicators be used?

Selecting indicators from lists or other projects requires consideration of the availability of data and the relevance of indicators to proposed targets. The tables below contain possible indicators with relevant data sources, but it must be emphasized that there is no one-size-fits-all indicator set for any purpose. In addition, multiple sources of data are encouraged in order to provide a more balanced understanding of the circumstances surrounding a data point.

³³UNDP Global Anti-corruption Initiative (GAIN) (2014-2017).

Table 9: Possible Indicators and Data Sources for Targeting Corruption³⁴

Target	Proposed Indicator	Results chain	Possible Data Sources	Type of Data	Global or National
Enact legislation designed to limit corruption	Increase the extent to which national laws are compliant with UNCAC	Input	UNCAC gap analysis and Self-Assessment Report	Administrative, Assessment	National, Some aspects can be globally compared
Reduce incidence of bribery cases by X % by year Y	Number of individuals that report paying a bribe when interacting with government officials Number of firms that report paying bribes to obtain services	Outcome	International Crime Victim's Survey TI Global Bribery Barometer Regional public opinion surveys Business Environment and Enterprise Survey (BEEPS)	Perceptions, Experiences	Global Regional National
Countries ensure adequate provision to detect and prevent illicit flows	Increases the extent to which national laws are compliant with UNCAC	Input	UNCAC gap analysis and Self-Assessment Report	Administrative, Assessment	National, Some aspects can be globally compared
Halve illicit financial flows by year Y	Volume of illicit financial flows	Outcome	Global Financial Integrity	Administrative	National, Global

Source: Iva Bozovic, UNDP Global Initiative on Anti-corruption (GAIN), 2014.

³⁴See section on Anti-corruption bodies for additional indicators.

Table 10: Possible Indicators and Data Sources for Improving Government Transparency³⁵

Target	Indicator	Results Chain	Possible Data Sources	Type of Data	Global or National
All national and local governments disclose information on budgeting, revenues and expenditures	Government budget data publicly available	Outcome	Agency reports, Compliance testing	Administrative	National
		Outcome	OGP Independent Reporting Mechanism, bi-annual assessment reports	Assessment, Administrative	National
	Quality and frequency of government budget data	Outcome	Open Budget Index Sub-scores	Assessment	National, Global
Legislation that provides for transparent corporate governance and accountability	Legislation exists for corporate reporting on social and environmental impacts	Input	National records; agency reports	Administrative	National
Enhanced state capacity regarding control of national resources	Transparent operations of extractive industries	Outcome	Extractive Industries Transparency Initiative (EITI)	Administrative	National, Global
Ensure open and transparent process for awarding public contracts	Existence of open and transparent process for awarding public contracts	Input	Vera Institute of Justice UN Global Compact World Bank Country Policy and Institutional Assessment	Assessment, Administrative	National, Global

Source: Iva Bozovic, UNDP Global Anti-corruption Initiative (GAIN), 2014.

³⁵See Annex 4 for additional transparency indicators at the national and organizational level.

Table 11: Possible Indicators and Data Sources for Enhancing Government Accountability³⁶

Target	Indicator	Results Chain	Possible Data Sources	Type of Data	Global or National
Hold officials accountable for misuse of office	"Whistleblower" legislation exists	Input	National records; agency reports	Administrative	National
			UNCAC gap analysis and Self-Assessment Report	Assessment, Administrative	National, Some aspects can be compared globally
	Procedures exist to hold officials accountable	Input	World Bank Country Policy and Institutional Assessment (CPIA)	Assessment	National, Global
	Percentage of court cases that cite corruption/bribery accusations	Outcome	Administrative records; agency reports	Administrative	National
All countries have legal instruments mandating income, asset, and conflict-of-interest disclosures by public officials	Public officials are obliged to file income, asset and conflict of interest disclosures	Output	UNCAC gap analysis and Self-Assessment Report, World Bank Public Accountability Mechanisms	Assessment, Administrative	National, Some aspects can be compared globally
	Citizens have the right to access disclosure records of members of the national legislature, civil servants, ministers, and executive	Output	Global Integrity Report, World Bank Public Accountability Mechanisms	Assessment, Administrative	National
	Citizens do access disclosure records of members of the national legislature, civil servants, ministers, and executive	Outcome	Global Integrity Report	Assessment	National
Increase citizen participation in providing input to policymakers	Participation mechanisms exist for informing policymakers	Input	National records; agency records	Administrative	National
	Percentage of major local government decisions in which input from participation mechanisms is taken into account	Outcome	Agency reports, monitoring local council decisions	Administrative, Assessment	National
	Percentage of people who say that they participated in a government-organized meeting or consultation	Outcome	Surveys such as Afrobarometer, Arab Barometer	Experiences	National, Regional
Increase citizen participation in policy-monitoring systems	Citizens are informed about place, date and topic of public discussion	Outcome	National records; agency records	Administrative	National
	Citizen satisfaction with provision of public services	Outcome	Citizen report cards, Focus groups	Perceptions, Experiences	National

Source: Iva Bozovic, UNDP Global Anti-corruption Initiative (GAIN), 2014.

³⁶See Annex 5 for additional accountability indicators at the national and organizational level.

Measurement Planning

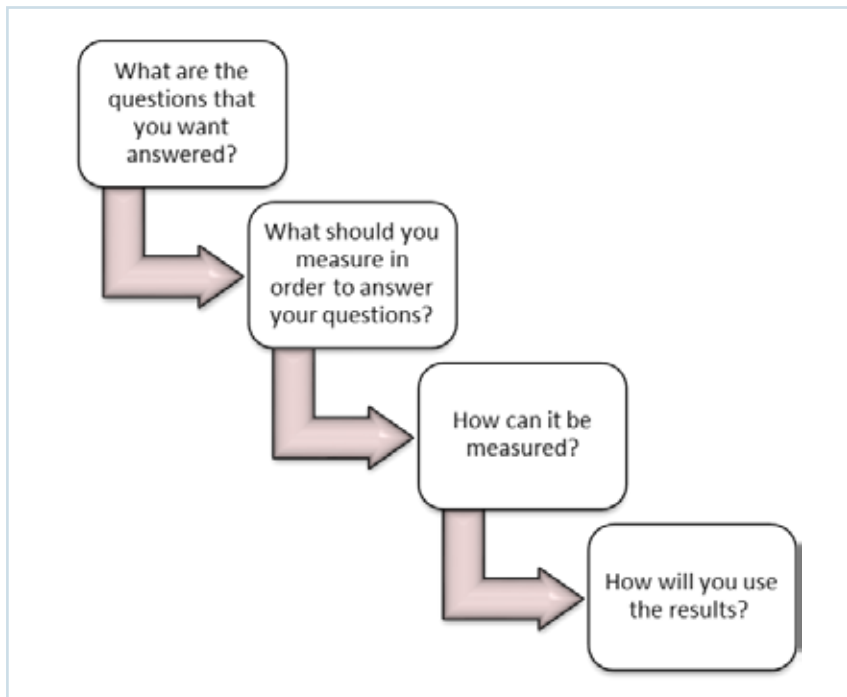
The most important question to ask when considering the use of measurement occurs before anything else – why do you need the data? The purpose of your data determines the most important parts of the project – its scope, stakeholders, results, and sustainability. The following questions are important for any measurement planning:

- Are you planning to use the data for policy reform, knowledge-building, or program monitoring?
- Is your priority to compare across countries, provinces, agencies, or communities? Or do you need data on one issue in one location?
- Are you trying to describe what is happening? Explain it? Do you want to follow the trail that the data leads you on, or test established theory?
- Are you trying to capture the outputs of agencies, projects, programs or interventions? Or are you trying to capture their impact on a wider domain?
- Are you trying to assess the levels, extent, or nature of actual corruption? Or are you trying to capture the existing risks (weaknesses, failures) in the system that might facilitate corrupt behavior?

None of these goals is exclusive. More than one purpose is certainly possible, but there are trade-offs. For example, cross-country comparisons often mean that detailed local data on one topic is not possible. As you dig deeper into country contexts, comparability becomes more difficult. Not every community has a formal justice system. Not every city has a publicly-operated sanitation service. Not every education ministry oversees the employment of schoolteachers.

Likewise, data intended to help change budget policies may not contribute much to a knowledge base located in a different part of the world, or to an expert in the field. It may assist public officials in making better choices about resource allocation in the community in which the data was collected. Once you decide the purpose of your data-collection effort, you can start asking questions about the research itself.

Figure 9: Four Strategic Questions to Guide Measurement Planning



Source: Author

What are the questions that you want answered?

Be sure to define your field of study very carefully.

What should you measure in order to answer your questions?

Unpack what you are trying to measure into discrete concepts. Consider the results chain – inputs, outputs, processes, outcomes – or the types of data – perceptions, experiences, assessments, or administrative data – that would best answer your questions. Too much data can be as problematic as too little data. Try to conserve resources and energy by designing clear indicators. Will the data tell you something that you need to know? Would the data help you make decisions about the intervention? Can you make those changes?

How can it be measured?

This step involves developing indicators and/or measurement approaches that will capture the relevant data. What kind of data will best suit your needs? Who is best positioned to collect the data, given credibility, data reliability, and purpose for the data?

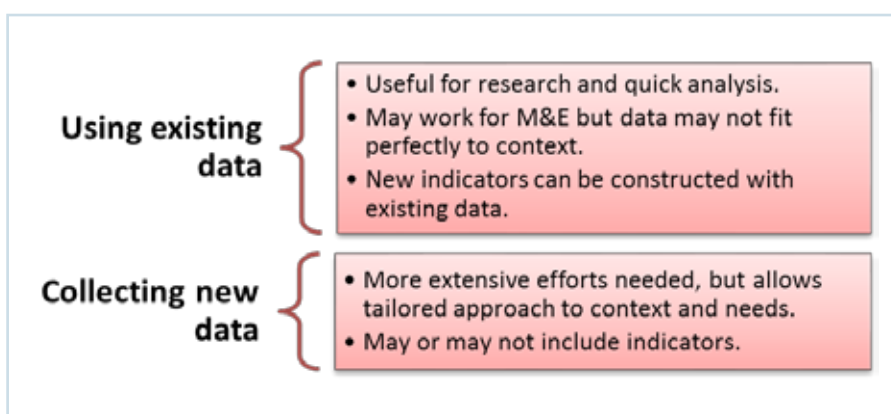
How will you use the results?

What are your follow-up plans once you have results? Which stakeholder groups will you involve in dissemination and what kinds of activities will best promote your data?

How Can It Be Measured? Finding the Data

Answering the third strategic question in the measurement planning stage can be done by either using existing data or collecting new data. The use of existing data involves identification of relevant datasets and selection of key indicators, or the combination or disaggregation of existing indicators. Collecting new data requires additional resources for data collection and quality control, and is discussed in more detail in the Methodologies section. Although it is much less costly to gather secondary data, it is more difficult to find or create appropriate measures with existing data that may have been collected for different purposes.

Figure 10: Finding Relevant Data to Meet Project Needs



Source: Author

There is a great deal of data that has already been collected, across countries, within countries, and across time, which can be used to achieve a variety of measurement purposes. There is not always the need to “mash-up” the data in complicated, statistically complex methods. Many external assessments may exclude the experiences of those groups most impacted by corruption: the poorest and most marginalized. It is possible to uncover the distinct experience of these marginalized communities by disaggregating survey data along many lines:

- Poverty
- Ethnicity/Language
- Gender
- Age group
- Geography: region, urban/rural
- Sector
- Health status
- Education level

Various methodologies already exist for constructing pro-poor and gender-sensitive indicators, such as household survey data, administrative data, and barometer data. However, most of the existing corruption and anti-corruption indicators hardly provide information on gender, poverty and other level of disaggregation. This limitation suggests collecting new data rather than using the existing data; however, collecting new data could be expensive and time-consuming. The quality of existing data, the objective of the usage of data and the cost of collecting new data are among major factors determining whether it is optimal to use the existing data or collect the new ones.

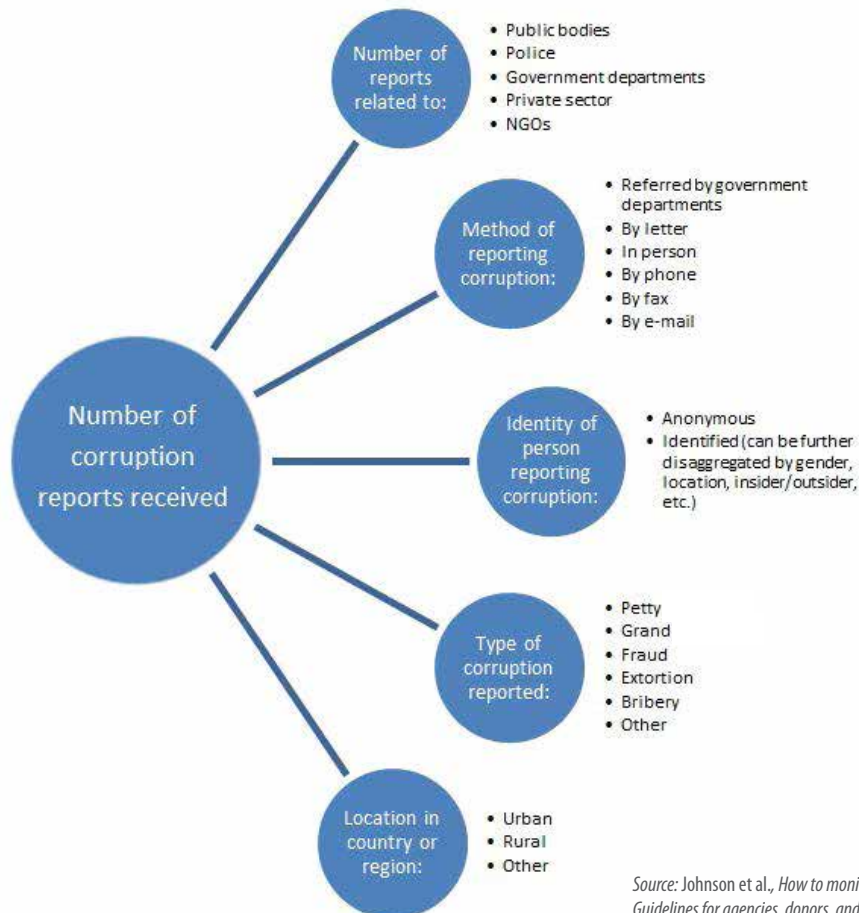
Table 12: Selected Pro-poor and Gender-sensitive Indicators for Corruption

Illustrative pro-poor indicators		Possible data sources
Poverty-status disaggregated	Percentage of poor households using public services who experienced corruption directly in the last 12 months	Survey
Specific to the poor	Percentage of reported corruption in public agencies of particular relevance to the poor, e.g., education, health, police	Survey
Implicitly pro-poor	Number of public agencies for which public expenditures tracking surveys are regularly conducted	Administrative
Chosen by poor	Percentage of poor households believing that corruption is unchanged or rising.	Survey
Illustrative gender-sensitive indicators		Possible data sources
Sex disaggregated	Ratio of women to men employed in civil service	Administrative
	Sex-disaggregated benefit incidence analysis of public spending on education and health	Gendered budget analysis
	Percentage of women in poor households having contact with administration in last 12 months who experienced corruption directly, compared with men in the same circumstance	Survey
Gender-specific	Existence of affirmative-action programs for women in the public ministries and administration	Survey
	Existence of anti-sex discrimination laws and equal opportunity policies in the civil service and evidence of enforcement and implementation	Analysis of legislation and implementation
	Annual expenditure on anti-sex discrimination legislation/equal opportunity policies	Administrative
Implicitly gendered	Incidence of reported corruption in the public agencies of particular relevance to women, e.g. access to safe drinking water, sanitation, and agriculture	Survey
Chosen by women	Level of satisfaction with public services expressed by women in poor households	Administrative, survey

Source: Adapted from Christopher Scott and Alexandra Wilde, *Measuring democratic governance: a framework for selecting pro-poor and gender sensitive indicators*, UNDP, 2006.

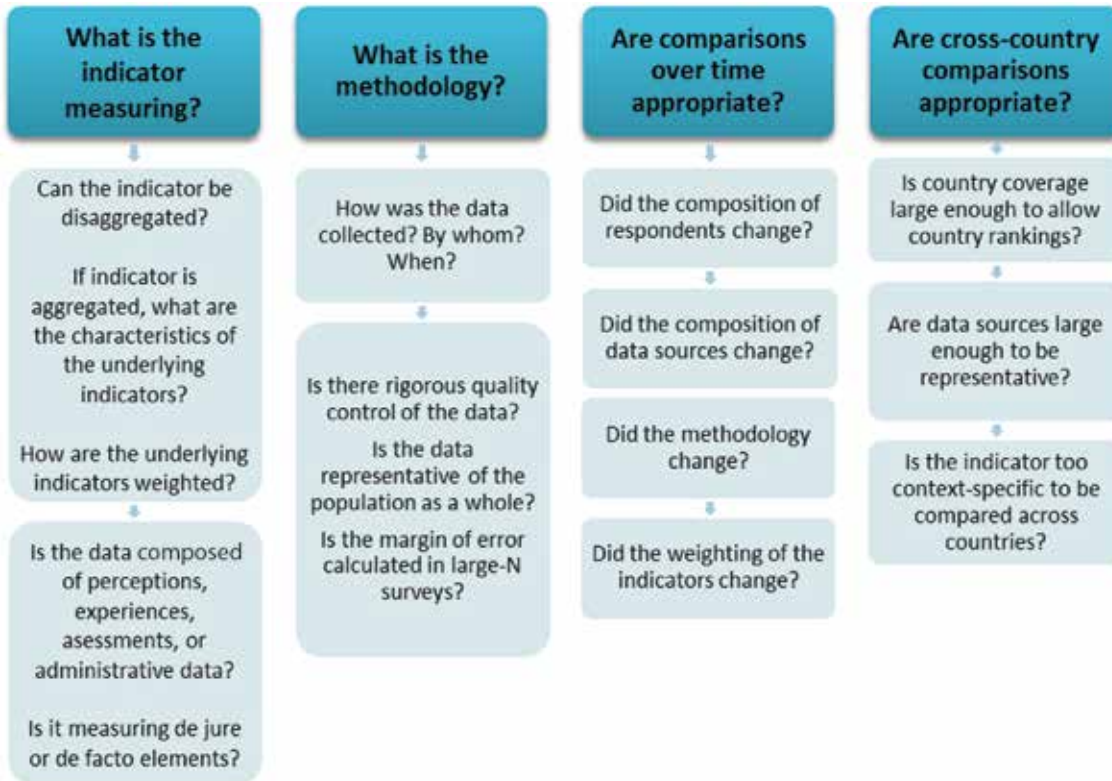
Data associated with individual corruption indicators can be disaggregated to capture data on the type of body which the complaint is made, the method of reporting, the type of corruption reported, as well as any data points about marginalized communities.

Figure 11: Disaggregation of Sample Corruption Indicator



When considering the use of existing data, it is important to understand the underlying indicators and methodology. It is important to determine the relevance of the indicator for a given aim, as well as understand its strengths and weaknesses. The questions in Figure 9 should be asked of any indicator being considered, to avoid inappropriate application and ensure fair interpretations are made from the data.

Figure 12: Considerations When Using Existing Data



Source: Author

Methodologies: Considerations and Constraints

The difference in definition between 'methods' and 'methodology' can be confusing, as they are used interchangeably to describe different things. Essentially, methods are tools used to collect data. A common example of a method is a survey, which can be employed in a variety of contexts. A methodology is the activity involved in meeting research goals, e.g., actions that will achieve the purpose established at the beginning of a measurement project.

For example, with the crowdsourcing method, data can be collected via the Internet on a number of different topics. But in order to meet project goals, this approach will need a methodology. This consists of activities that shape the project from beginning to end, always building toward the purpose of the project. Below is an example of a methodology employing crowdsourcing that aims to change government policy affecting a specific group of people:

- (1) engage with stakeholders to determine the best parameters for the project (e.g. timing, topics, indicators, hosting services, language, etc.)
- (2) build a user-friendly website with those parameters
- (3) ensure that the site is advertised to get the most exposure to your target group of respondents
- (4) identify ways to validate the data, e.g., how will you determine whether data is about bad management or corrupt behavior?
- (5) aggregate or analyze the data in a way that answers your research questions
- (6) disseminate the results to the appropriate audiences, e.g., media, NGOs, donors, government officials, community members, all of whom will respond in different ways
- (7) engage with government officials and community stakeholders to reform policies or procedures.

There are many kinds of methodologies that can be built around the method of crowdsourcing, or any other method, or any combination of methods. Before you select or adapt a methodology, however, you must know what you want to measure and the purpose of your measurement.

Once you have preliminary answers to these questions, consider the various methodologies that have been used for similar projects. The table below contains common approaches to measuring corruption, transparency, accountability, and integrity, along with important factors to consider when designing or adapting a methodology.

Table 13: Selected Methodologies, with Factors to Consider when Adapting to Different Contexts

	Description	Use/Purpose	Data collection	Expertise	Stakeholder participation
Citizen Report Card	Surveys that provide a quantitative measure of user feedback on the quality, efficiency and adequacy of different public services	Used in situations where demand side data is absent, including user feedback on quality and satisfaction with public services, and instances of bribery or petty corruption. Outputs may be used as basis for performance based budget allocations, cross-state comparisons on quality of public services, and supplementing national service delivery surveys.	Representative survey of users (formal stratified random sampling) at city, state, or national level	Technical competence to scientifically design, execute, and analyze the survey	(a) in the design of questionnaires where the performance indicators and key issues are developed through focus group discussions with citizens, (b) during the survey execution, where qualitative interviews are used to support questionnaire data, and (c) during dissemination where a variety of NGOs are brought in to use the data for advocacy and reform.
Social Audit	Participatory auditing process that collects information on the resources and performance of an organization	The central objective of a social audit is to monitor, track, analyze, and evaluate government performance. The information is analyzed and shared publicly in a participatory, constructive fashion.	Social Audits use a combination of different methods for obtaining the relevant data including interviews, surveys, quality tests, compilation of statistics, case studies, participant observation, evaluation panels, gathering relevant official records and extracting relevant information from existing data of various sources.	Technical competence in the design and execution of a variety of data-collection methods and analysis of differing types of data	Social auditing is most effective when government authorities are willing to collaborate and allow civil society to provide feedback and information about their activities. The stakeholders should be identified and included in the whole process. The stakeholders should be a mix of government actors from different levels, service providers and/or contractors, representatives of civil society organizations, beneficiaries, and workers of the service providers or contractors.
Community Score Card	Community-based monitoring tool that is a hybrid of the techniques of Social Audit, public expenditure tracking, and citizen report cards	Used in community monitoring of public-sector agency at the local/facility levels, combined with an interface meeting between users and providers to provide respective feedback and generate a mutually agreed reform agenda. Facilitates community empowerment and agency transparency.	Focus groups in communities at village level: (1) citizens and (2) service providers	Competence to facilitate the focus groups and the interface meeting	(i) identifying local leaders to assist and participate, (ii) developing community performance scorecard with community focus groups, (iv) generation of the self-evaluation score card with facility/project staff focus groups, (v) the interface meeting between community and providers
Risk Assessment (See Transparency International "Corruption Risk Assessment: Topic Guide")	A (diagnostic) tool to identify weaknesses within a system that may facilitate corrupt practices. It focuses on the potential for – rather than the perception, existence or extent of – corruption.	Used in institutional and sector assessments to strengthen organizational practices and monitoring arrangements.	Combination of secondary sources (legal-institutional analysis and desk research) and primary sources (household surveys and questionnaires, focus groups, key informant interviews, checklists).	Technical competence in the design and execution of a variety of data-collection methods and analysis of differing types of data	Risk assessments have not historically been participatory mechanisms, but they benefit from the insight of employees and managers, as well as citizens on the receiving end of government services. These data are collected through household surveys and questionnaires, focus groups, and key informant interviews.

<p>Integrity Assessment</p>	<p>Similar to risk assessments in that they aim to identify vulnerabilities to corruption and actual levels of corruption. But they go further to examine organizational ethics and administrative culture, which involves study of the values and behaviors of public officials, and the constraining rules that attempt to mitigate risks or conflicts of interest.</p>	<p>Used in institutional and sector assessments to strengthen organizational practices and monitoring arrangements.</p>	<p>Combination of secondary sources (legal-institutional analysis and desk research) and primary sources (household surveys and questionnaires, focus groups, key informant interviews, checklists).</p>	<p>Technical competence in the design and execution of a variety of data collection methods and analysis of differing types of data</p>	<p>Integrity assessments benefit from the insight of employees and managers. These data are collected through questionnaires, focus groups, and key informant interviews.</p>
<p>Public expenditure tracking surveys</p>	<p>Diagnostic or monitoring tool to understand problems in budget execution. The focus is to identify delays and unpredictability of public funding, leakage and shortfalls in public funding, abuse of discretion in resource allocation</p>	<p>Used in community-level contexts to track flows of budget from center of government to local levels.</p>	<p>Representative survey of users at different levels of government (formal stratified random sampling), and administrative data.</p>	<p>Technical competence in the design and execution of a variety of data-collection methods and analysis of differing types of data</p>	<p>Civil society groups often initiate and conduct PETS, sometimes with the help of donor expertise. Design of surveys should involve insight from citizen groups to identify variables of interest. Interaction with government officials and user groups is important in the dissemination and follow-up phases.</p>
<p>Exit/User Survey</p>	<p>Surveys that are conducted immediately after users obtain a service</p>	<p>Used to collect short user feedback on quality and satisfaction with public services, and instances of bribery or petty corruption.</p>	<p>Brief in-person interviews or online surveys</p>	<p>Minimal technical competence to design, execute, and analyze the survey</p>	<p>Surveys should involve citizen groups in the identification of variables of interest, and again at the dissemination phase. Interaction with government officials is important to ensure accountability mechanisms will be triggered.</p>
<p>Citizen/Public Official Survey</p>	<p>Surveys that collect information on the experiences, satisfaction, knowledge or crime victimization of individuals. They may also be used to collect information from public officials or government representatives, often referred to as self-reporting.</p>	<p>Used to collect comprehensive information on experiences or victimization, often to supplement administrative data or national service delivery surveys. Also used in lieu of administrative data when it is unavailable.</p>	<p>Representative survey of users (formal stratified random sampling) at city, state, or national level. Telephone or in-person survey questionnaires</p>	<p>Technical competence to scientifically design, execute, and analyze the survey</p>	<p>Surveys should involve citizen groups in the identification of variables of interest, and again at the dissemination phase. Interaction with government officials is essential to ensure accountability mechanisms will be triggered.</p>
<p>Household Survey</p>	<p>Surveys that are administered at the household level. Similar to a citizen survey, but often has many more questions, and data is collected through face-to-face interviews.</p>	<p>Use to collect information about the household and the individuals living in those households. Household surveys are usually administered regularly by national statistics offices and cover a wide range of topics, using the household as a unit of analysis (as opposed to individual citizens, users, or officials).</p>	<p>Representative survey of households (formal stratified random sampling) at city, state, or national level</p>	<p>Technical competence to scientifically design, execute, and analyze the survey.</p>	<p>NGO involvement in the development of household surveys is becoming much more common, as local community members and organizations can provide critical insight into issues of governance and corruption. One goal of SHASA³⁷ is to bring together national statistics offices and local NGOs to improve the collection of this type of data in household surveys.</p>

³⁷Strategy for the Harmonization of Statistics in Africa, led by the African Union Commission, African Development Bank, and UN Economic Commission for Africa.

Table 13: Selected Methodologies, with Factors to Consider when Adapting to Different Contexts (continued)

	Description	Use/Purpose	Data collection	Expertise	Stakeholder participation
Expert Survey	Expert assessments on specific areas of corruption or anti-corruption, usually global or regional in scope	Used to capture data about difficult-to-measure topics, such as when local data or administrative data is not available, or is not comparable across countries.	Non-representative survey of experts (1-5 per country) with peer/expert review serving as a reliability check.	Technical competence to scientifically design, execute, and analyze the survey. If a weighted index is generated, knowledge of aggregation methods is essential, particularly margins of error.	Stakeholders in expert surveys are often employed to establish the validity of the indicators (e.g., is there enough evidence to assess the indicator at all?) and reliability of the data (do all experts who assessed the indicator agree on the results?). Community and NGO stakeholders should also be involved in project design to ensure that the purpose of the project is adequately reflected in the indicator development. For example, asking about topics that are not relevant for user communities will make it more difficult to encourage demand-side policy reform.
Crowdsourcing	Practice of collecting data by soliciting contributions from a large group of people, and especially from an online community, rather than from traditional surveys	Used to aggregate many voices on a topic that may be missed out in traditional surveys, especially concerning collection of real-time data.	Online or SMS platform. Validation of submission is an ongoing challenge.	Technical competence to design website collection and analytical methods	Local stakeholders are essential to the successful launch of crowdsourcing platforms for anti-corruption purposes. Because it is a demand-side approach, information dissemination is crucial to enlisting the participation of respondents. Also important is the use of data for reform purposes, which serves as an incentive for participation. NGOs and community organizations can facilitate all of these components.
Compliance/ Field Testing	Testing to determine whether a service complies with standards and laws.	Used by citizen groups to investigate the quality of delivery of promised government services	Engaging with the service to test whether it functions. May involve site visits and applications to government offices.	Competence to design testing strategies and analyze results.	Citizen groups and individual citizens are essential to ensuring the success of compliance testing, as they are tasked with site visits and government interaction.

Measurement Design Considerations

Once a theory of change has been established, and targets and indicators have been identified, designing the related measurement project requires consideration of key factors. Stakeholders play many roles in measurement design and implementation, and where appropriate, should be considered equal partners in measurement approaches. In some cases, capacity building can occur simultaneously with project implementation. This is an effective means of instilling the norms of monitoring and evaluation into organizations. Quality control is essential in any measurement approach, and contributes, among other things, to the sustainability of data-collection efforts. In fact, establishing credibility is often achieved through rigorous quality control and/or meaningful stakeholder involvement. This is because no method can be termed as standard unless it is agreed upon by major stakeholders. Similarly, no measure can be constructed that provides a perfect reflection of reality given the multi-faceted nature of corruption and anti-corruption.

Stakeholders

A major consideration in project design is the role of stakeholders. Stakeholders include governments and public officials, community members, users of the data, as well as donors and experts. Depending on the nature of the project, various stakeholders can be involved to provide insight, oversight, and much-needed credibility of the resulting data. Regardless of the potential complexity in the research design, aggregation methods, or data-collection process, stakeholders are a crucial element to the success of a project. Various kinds of stakeholders can be included at all stages in the design process: project design, indicator development, data collection, data analysis, and dissemination of the data.

Service users and community members are the most overlooked stakeholders, even though they are often the group that is most affected by corruption and lapses in transparency, accountability, and integrity. While they may not have research expertise, they hold valuable information and experience about many of the ways that government activities fail to meet standards. If you are aiming to hold agencies accountable for service delivery challenges, it is a good idea to consult with the community that is most affected. Users and community members can provide insight into the types of problems they regularly encounter, assist with the data-collection efforts. In some social accountability methodologies, they are an integral part of holding government to account with the resulting data. Advocacy efforts are an important use of data in demand-driven campaigns for reform.

Non-governmental Organizations (NGO) often play the role of project implementer for donors, retained in a limited or supporting capacity. But NGOs are far too valuable a resource to be excluded as priority stakeholders. In cases where NGOs are well connected to their communities, and possibly well regarded by public officials, they might be considered “political entrepreneurs, with the credibility and capacities to build mutual trust and relationships among diverse actors.”³⁸ But NGOs may need capacity building to participate as collaborators or initiators of extensive measurement projects.

³⁸Brendan Halloran, *Thinking and Working Politically in the Transparency and Accountability Field*, Transparency and Accountability Initiative, 2014.

Participatory Monitoring and Evaluation

M&E is often facilitated by donors, for donors, as a means of establishing upwards accountability and informing decision-making processes for similar projects. But measurement processes should also promote several key elements: local ownership through participatory mechanisms if appropriate, capacity development of national systems, and inclusiveness of marginalized groups. Participatory monitoring and evaluation (PM&E) is an approach that involves local stakeholders in various stages of design, data collection, and analysis. PM&E takes local knowledge into account and facilitates local ownership of the process, allowing affected individuals and communities to participate and benefit from the information gathered. By employing local knowledge, PM&E improves project effectiveness and verification of data. It also complements project monitoring in decentralized, or isolated areas, and where state agencies lack monitoring capacity.³⁹ But there are limitations to PM&E. In particular, it is labor-intensive, costly, and may not produce as reliable or accurate data as a donor-driven M&E project.

Source: Author

Governments and public officials hold the key to transformation of poorly administered processes and services. Left out of the process, they are less likely to believe in the project findings and more resistant to change. Governments are clearly important end-users, as the purpose of policy-relevant data is to spur improvements in public-sector and governance processes. It is important to understand that reluctance to participate on the part of government officials may not extend to all stages of project design. Keeping channels of communication open throughout the process allows for inclusion of officials in later stages. Confrontational or contentious interaction with government officials (as opposed to collaboration) is sometimes necessary, but not necessarily the first or best option.

Donors are clearly resources for assistance with stakeholder coordination, funding, and dissemination. They can also be important for establishing credibility with governments and communities, and for establishing channels of communication with public officials and other organizations working on the same topics.

Experts can provide advice on all stages of the project design, and their participation at key points in the project may keep it on track when faced with difficulties. In the design stage, it is important to ensure that indicators are capturing the right kind of data to answer your questions. Experts can shed light on the best types of data and the most efficient means of capturing that data so that the project can be realized.

Consideration of **data users** should not be left until the end of the measurement project. Project designs can be tailored to capture data that is relevant for various actors, as different groups of individuals may utilize data for a range of purposes. NGOs use data both to confront and to collaborate with public officials. Journalists use data to prompt discussion of issues such as corruption, service delivery, and access to information. Government officials use data to improve the administrative processes within their units, and also to implement safeguards to prevent corrupt behavior. The private sector uses data in a variety of forms to determine areas for investment and transnational business activities, which is an often-overlooked, but powerful, use of governance and public sector data.

³⁹Warren A. Van Wicklin III and Asli Gurkan, *How-to notes: participatory and third party monitoring in World Bank projects - what can non-state actors do?*, World Bank, 2013.

De Jure vs. De Facto Measurements

The **de jure institutional framework** consists of the formal rules governing the actions of individuals or organizations. These rules consist of *laws, policies, operating procedures, and/or administrative regulations* that assign responsibilities and authority to act.

Relationship to outcomes: *All of these aspects of a system are integral to outcomes, as organizations need rules and procedures to achieve targets. Even though the precise relationship between the de jure framework and outcomes is not clear, it is important to establish clear rules and regulations that signal government commitment to anti-corruption goals.*

Organizational capacities are the resources employed by the individuals or organizations in order to achieve their goals. These capacities include the existence, amount, or number of particular resources, such as money, personnel, equipment, facilities/buildings, etc. They also include the quality of those resources, including age of technology, qualifications of staff, quality of records and information management, etc.

Relationship to outcomes: *Organizational capacities are the underlying components that drive outputs and outcomes, and their improvement may result in substantial reform progress. However, they are often overlooked in measurement approaches. The capacity assessment of any organization should recognize the variety of political, cultural, legal and administrative circumstances in which various organizations such as anti-corruption agencies operate.⁴⁰*

The term **de facto** refers to the implementation of the de jure framework. In measurement terms, however, it can refer to a number of different elements along the results chain.

Relationship to outcomes: *De facto could include short-term outcomes like establishment of offices. Or it may capture intermediate outcomes that reflect changes in behavior, such as increased rates of grievance/complaints redress, sustainable funding for oversight institutions, or increased civil society engagement in participatory monitoring. It could also refer to long-term outcomes (or impacts) like improvements in service delivery or decreased crime and victim rates due to the improvement in performance resulting from the changed behavior of the institutions or actors.*

But impacts are extremely difficult to capture efficiently, and more importantly, with accuracy. They often do not emerge for a long time. Impacts of corruption and anti-corruption are likely the result of a complex interplay of factors, and as a result, they might be difficult to predict or to attribute to certain factors. It is important to remember that perceptions of the impact of corruption are not the same thing as the actual level or extent of corruption. This precise kind of data is hard to generate, because corrupt activities are covert by nature, making identification nearly impossible without reliable administrative data and sophisticated methods of detection and analysis.

⁴⁰UNDP, *Practitioners' Guide: Capacity Assessment of Anti-Corruption Agencies*, 2011.

For the measurement of corruption, de facto often means the “impact of corruption on quality of life and public service delivery.”⁴¹ This can be interpreted as the experiences of users or citizens and the performance of government functions. These two aspects of de facto realities complement one another: Experiences reflect demand-side problems, and performance captures supply-side challenges. Performance includes a variety of activities, such as the delivery of services, enforcement of rules or laws, production of key budget documents, etc. Performance is a measure of how well the organization is meeting its goals or targets, which is another way of referring to how well the organization is implementing the laws, or whether the organization is complying with the laws at all. Good performance happens in spaces of transparency, accountability, and integrity. Thus, at the sector, institutional, and community level, de facto performance measurement can be used to measure outcomes of the extent and nature of transparency, accountability, and integrity within government (or the private sector).

What is the Implementation Gap?

In general, the implementation gap between de jure and de facto refers to the “difference between the country’s legal framework surrounding good governance and anti-corruption and the actual implementation and enforcement of that same legal framework.”⁴² Measuring both de jure and de facto aspects allows for a comparison of policy adoption and policy outcomes. Government commitment to reform can be assessed regularly in a number of areas, allowing comparisons across sectors and countries. This is particularly relevant in the case of global conventions or initiatives that a large number of countries have ratified. But it is also important to understand the implementation gap at the organizational or agency level, where service delivery happens. Communities often have limited access to centers of decision-making, but bear the consequences of poor implementation most heavily.

Quality Control

Quality control of data is important to establishing credibility with stakeholders. Without reliable, valid, accurate data, there is little chance that data will be easily disseminated or used for the purposes that it was intended. There are recognized methods for establishing validity and reliability with survey data, particularly with calculating margins of error.⁴³ There are also methods for establishing reliability with qualitative data or assessments that involve peer-review processes. Triangulation is an alternative approach where multiple sources and types of data are used to measure the same phenomenon. Similarly, peer reviews by experts are often used to control the quality and integrity of the data-gathering processes.

⁴¹Nadgrodkiewicz et al., *Improving Public Governance: Closing the Implementation Gap Between Law and Practice*, Center for International Private Enterprise (CIPE) and Global Integrity, 2012.

⁴²Nadgrodkiewicz et al., 2012.

⁴³Examples include: test-retest reliability, parallel forms reliability, internal consistency reliability (e.g., average inter-item correlation, split-half reliability), criterion-related validity, formative validity, sampling validity, predictive validity, concurrent validity, construct validity, convergent validity, discriminant validity, consequential validity. Further reading on these topics is available in statistics texts.

Validity

The validity of data is a measure of how closely it is related to the phenomenon being studied, which is often a subjective judgment by experts or a group of stakeholders. One common issue in corruption measurement has been the validity of corruption perceptions for measuring actual corruption. In other words, are perceptions of corruption evidence of the extent of corruption in a country? No. Perceptions data reflects the opinions of individuals, and it can rise and fall independently of the actual levels or extent of corruption. Perceptions are not evidence of corruption; they are evidence of how people feel or what people think about corruption. Therefore, perceptions are not necessarily considered a valid measure of actual corruption. Since perceptions are often based on experience, however, if a large percentage of the study sample believes that corruption is a problem in their country, it is likely to be true. But this says little about the nature or amount of corruption that exists.

Perceptions of corruption are considered a proxy indicator of corruption. Measuring the actual levels of corruption, or the extent of corrupt practices, is extraordinarily difficult. Many datasets employ proxy indicators that stand in for more direct, but unattainable, measures. Proxy indicators tend to be very context-specific, because different contexts have varying traditions, social structures, and societal values. Proxy indicators for petty corruption may include the number of steps needed to obtain a construction permit, the number of traffic stops by law enforcement, or the inexplicable delay in processing times for court cases involving similar crimes. But care must be taken not to equate the proxy indicator exactly with the underlying phenomenon. They are different, but related, measures. Petty corruption, such as bribery and fraud, can sometimes be directly measured with reasonable accuracy, but in other cases, and with other types of corruption, including patronage, conflict of interest, abuse of power, exploitation, etc., corruption escapes direct measurement.

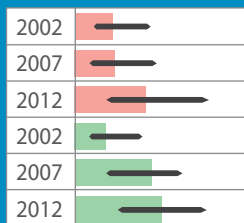
Reliability

No data point is free from potential bias or error. **Reliability** is the extent to which we can rely on the accuracy of the data. In fact, consistency is the main measure of reliability.⁴⁴ Survey questions can be tested for reliability with a number of design methods and statistical tests. With qualitative data, inter-rater reliability measures are employed. **Inter-rater reliability** is used to evaluate the degree to which different judges or “raters” agree in their assessment decisions. Inter-rater reliability is useful because people will not necessarily interpret indicators the same way; raters may disagree as to how well certain responses or material accurately capture the topic being assessed. A modified version of inter-rater reliability used in expert surveys is called “peer or expert review.”

⁴⁴Roger Pierce, *Research Methods in Politics*, 2008.

What Are Margins of Error?

The margin of error helps you estimate how close you are to describing something about a population based on your sample data. Surveys can't cover the entire population of a community or a city. But practitioners can use techniques to create a sample of the population that is representative. This means they can make claims about the population without having to get responses from everyone. Margins of error explain how much the sample data differs from the responses of the entire population at the point of time when data is collected. Consider this: A survey reports that 45% of the *sample* paid a bribe last year, with a margin of error +/- 3%. This means that between 42% and 48% of the *population* would report paying a bribe. Margins of error are thus very important when comparing data.



Consider this: Aggregated indices typically report margins of error with thin lines through the graph columns. These margins of error reflect the extent of agreement among the underlying data sources: When data sources tend to agree, the margins of error are smaller, and when they disagree, margins of error are larger. For this reason, lower margin of error requires a larger sample size.⁴⁵ When the margins of error overlap, there is a great possibility that straightforward comparisons between scores are not possible. None of the scores in the example can be meaningfully compared, as all the margins of error overlap. You can get a general idea of the score across countries or years, but you cannot compare those scores with much confidence.

Source: Author

Triangulation

Triangulation is an alternative method for establishing validity and checking reliability in measurement projects by analyzing a research question from multiple perspectives.

- *Methodological triangulation* involves using more than one method to gather data, such as interviews, observations, surveys, crowdsourcing, and/or desk research.
- *Data triangulation* involves using different sources of information and different types of data. For example, surveys can be sent to citizens, public officials, and/or households, and data collected can range from perceptions to administrative data.

Triangulation of research methods is often necessary in order to draw meaningful conclusions from the data, particularly when concepts are fuzzy or data is difficult to collect. It involves the validation of data through more than two sources, facilitating a better understanding of the phenomenon being studied.

⁴⁵Higher confidence levels also require a larger sample size. The margin of error is the amount of error that can be tolerated, while the confidence level is the amount of uncertainty that can be tolerated.

Sustainability

The sustainability of data-collection efforts is related to a number of project components, and not just the complexity of the project design. These factors include the number of countries covered and the regularity with which data is collected. Sustainability is also influenced by the level of analysis, the specificity of the indicators, and the conceptual fuzziness of the object of study.

Attempts to embed methodologies and indicators into the existing administrative systems of governments are an excellent way to ensure sustainability. Examples include using indicators of performance assessment in the human resources management systems, or implementing policy-relevant indicators in departmental or sector monitoring and evaluation systems in the public sector. Embedding new indicators into already existing systems does not only encourage sustainability; it also allows public-sector managers to claim credit for administrative successes, and demonstrates to civil society that engagement matters.

Constraints

There are strengths and weaknesses to any approach to measurement, just as there are risks and limitations. When trying to decide the parameters of a data collection project, you should consider some of the most important constraints. In many cases, you can better prepare for constraints by conducting a small pilot exercise to test your project design. Pilot projects are helpful for estimating the possibilities for full-scale project, and also to determine if your project design will lead to results that answer the research questions.

Costs are dependent on several interrelated factors. Data-collection costs tend to dominate the budgets of many measurement projects, because they are the most time- and resource-intensive stage. Ongoing projects that require data collection every 6 months or every year will be more expensive than projects that collect data every 2-5 years. Projects with detailed survey questionnaires and several thousand respondents may be more expensive than short expert surveys. However, if the expert surveys are global in scope with complex indicators, they might end up being more costly than one-country large-N surveys with short, easy-to-answer questionnaires. Labor costs include project leaders, quality-control staff, data analysts, and staff that engage in dissemination activities. In expert survey projects, country researchers might be paid for the assessment participation, but in large-N citizen surveys, respondents are often not paid at all.

Expertise may play a role in determining which methodologies can be used effectively. It is quite difficult to learn the skills on-the-job without some research or prior experience. Simple surveys with few respondents may not need statistical expertise. But large-N surveys, or aggregated indices, require familiarity with statistical methods to eliminate bias and calculate margins of error. Quality control over data collection with large numbers of indicators may not be possible without clear procedures and oversight that an experienced researcher can provide. Indicator-driven case studies, audits, and risk assessments require familiarity with several qualitative research methods such as interviews, focus groups, document review, and compliance testing.

Political will/enabling environment: The context in which measurement initiatives are conducted determines much about the approach and methods involved. But the larger context of power dynamics among agents, structures, and processes will influence the impact and dissemination of the data. Before planning to produce or use data, practitioners must assess the civic space available for measurement and accountability, and adjust expectations accordingly. In particular, the risks of engagement by local partners should be considered a major constraint when developing measurement strategies.

Challenging Impunity is Most Risky for those with the Least Clout

Tactical information interventions are often based on the implicit assumption that participation has more benefits than costs, if the costs are recognized at all – and that the people who are targeted for encouraging participation perceive the benefits as being greater than the costs. That's one reason that the role of external allies is so important, insofar as they might be able to reduce the risks inherent in challenging impunity from below, as well as their capacity to help to identify actionable pathways through which collective action could leverage a response from power-holders. That's the substantive meaning behind the technocratic-sounding term “enabling environment.”

This story is about a recent incident involving a Mexican NGO partner working in the indigenous highlands of Chiapas, where they trained community members to become local accountability promoters. Local citizens learned how to exercise their legal rights to call on their mayors to disclose how public funds were spent, in socially sensitive workshops in their own language. Everyone knew that most of those mayors are local bosses who play rough.

The local NGO reported that yes, we know it's risky, but we think we have enough momentum and community support to pull this off. So they drew on their long-standing social networks, trained dozens of local activists and launched their bottom-up, public campaign for accountable local government. The backlash came hard and fast. The NGO was banned from the district, and they headed back to their headquarters in a regional city. But the consequences were not so simple for the local organizers, at least one of whom was expelled from his own village, under threat of violence. Local radio media coverage of this scandal was good and national news coverage was nice, but the local mayors didn't care. The power of shame didn't work on the shameless. A flagship national NGO blogged about it, but their coverage stressed how great it was that the grassroots campaign was launched, mentioning only in passing that the local community watchdogs got screwed – a reflection of our community's persistent tendency to find the glass at least half-full, even when it's almost empty. The very worried NGO then met with state government, which promised to do “trainings” for local mayors to raise their awareness about the right to information, but they didn't show up to their own workshops. The institutional reforms of the national information access regime then being debated in Mexico City might as well have been happening on another planet. The mayors retained their impunity.

Source: Jonathan Fox, Seven tensions facing the transparency/accountability agenda, 2014.

Evaluation and Impact Assessment

This chapter presents various approaches to the evaluation of effectiveness of anti-corruption projects, programs, and activities. Both experimental impact evaluation and theory-based evaluations are discussed. The second half of the chapter presents different examples of measuring effectiveness at the project/sector level, institutional/organizational level, and national/international level.

“There is the question of measuring corruption, and there is the question of measuring the effectiveness of anti-corruption interventions. They are often conflated, but they are in fact, two very different things.”

- Jesper Johnson, U4 Chr. Michelson Institute, Norway

Impact evaluation is an assessment of completed activities in order to attribute causality or determine the extent of contribution to external outcomes. Evaluation of impact in the field of corruption and anti-corruption is aimed at assessing the effectiveness of an activity or project at reducing corruption or enhancing transparency, accountability, and integrity.

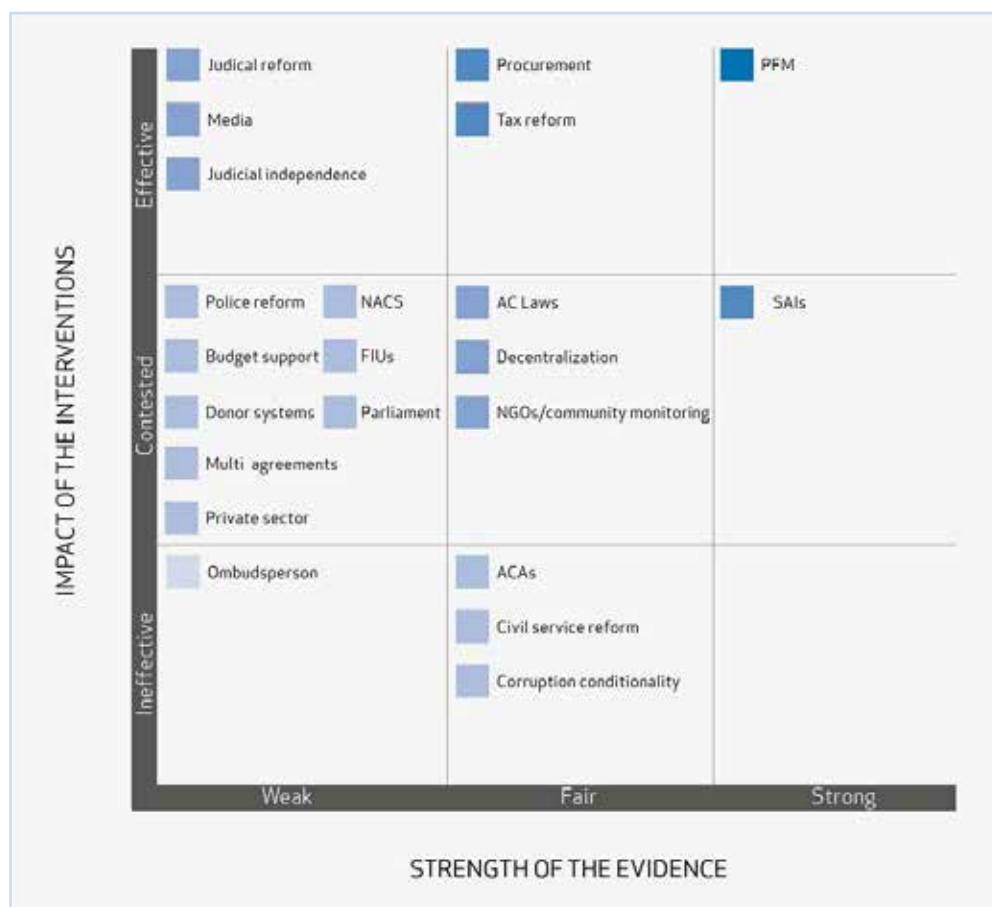
An evaluation is an assessment, as systematic and impartial as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area, institutional performance, etc. It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors of causality, in order to understand achievements or the lack thereof.⁴⁶

There is a considerable lack of knowledge about the effectiveness of anti-corruption interventions due to weak reporting and evaluation standards. However, there is some evidence that interventions in particular sectors are more effective than others, including public financial management, procurement, and tax reform. The lack of strong evidence for other areas of focus is not necessarily due to failure of policy, but rather, can be attributed to a failure to effectively monitor progress and evaluate impact.⁴⁷

⁴⁶UNDG, *Results-Based Management Handbook*, 2011 (page 34).

⁴⁷Johnson et al., *Mapping evidence gaps in anti-corruption: Assessing the state of the operationally relevant evidence on donors' actions and approaches to reducing corruption*, Chr. Michelsen Institute (CMI), 2012

Figure 13: Summary of the Evidence from a Systematic Review of Anti-corruption Interventions



Source: Johnson et al., *Mapping evidence gaps in anti-corruption: Assessing the state of the operationally relevant evidence on donors' actions and approaches to reducing corruption*, Chr. Michelsen Institute (CMI), 2012.

Experimental Impact Evaluations

The purpose of an experimental impact evaluation is to determine whether a project has achieved its intended impact, and more specifically, to quantify the size of that impact. It aims to establish causality between intervention and impact. It is often referred to as counterfactual analysis because it estimates outcomes in the absence of the intervention, and impact is estimated by comparing counterfactual outcomes to those observed under the intervention.⁴⁸

These randomized, controlled trials, also referred to as experimental and quasi-experimental impact evaluations, have been applied in corruption and governance interventions, but the findings have been mixed. This is likely because impact evaluations presume a short-term, linear results change that is quite rigid. Governance and corruption interventions are complex, with diverse factors contributing to outcomes, as dictated by political economy concerns (e.g., stakeholders, credibility concerns) and the enabling environment (e.g., elections, violence, macro-economic stability). This type of evaluation may be poorly suited to the non-linear, complicated change processes that characterize transparency and accountability interventions.⁴⁹

⁴⁸See Annex 2 for descriptions of the different types of quasi-experimental and experimental methods.

⁴⁹For more detailed explanation of the weaknesses in randomized controlled trials (RCTs) in the area of human affairs, see Michael Scriven, "A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research," *Journal of MultiDisciplinary Evaluation*, 2008.

Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. (Ferraz & Finan, 2008)

Researchers in Brazil sought to investigate whether making audit information on political corruption publically available would affect the electoral outcome of incumbent mayors. The randomized assignment provided an opportunity to observe whether voter-access to information about a politician's corruption level prior to the election impacted the average vote share and re-election rate for incumbent mayors.

For every additional corrupt violation reported against an official, releasing the audit results reduced the likelihood of re-electing that official by approximately 20 percent. The effect of the policy was similar for other measures of electoral performance, such as the change in vote share and margin of victory. These results suggest that voters not only care about corruption, but once empowered with the information, they update their prior beliefs and punish corrupt politicians at the polls.

In municipalities with local radio stations, the effect of disclosing corruption on the incumbent's likelihood of re-election was more substantial. Results indicated that for municipalities that released audit results prior to the election and revealed at least one count of corruption, the presence of an additional radio station decreased the incumbent's probability of re-election by 10.7 percent. Not only did radio stations increase the effect of the audit when corruption was revealed, they also promoted the re-election of non-corrupt incumbents. When corruption was not found in a municipality with a local radio station, the audit increased the likelihood that the mayor was re-elected by as much as 20 percentage points.

These results indicate that the disclosure of information enhances political accountability in the very specific context of voting.

Source: Poverty Action Lab, <http://www.povertyactionlab.org/>

The GoBifo Project Evaluation Report: Assessing the Impacts of Community-Driven Development in Sierra Leone (Casey, Glennerster, Miguel, 2013)

To both prevent a return to violence and to stimulate economic development, the Government of Sierra Leone implemented a number of reforms that gave communities, and vulnerable groups within them, a greater voice in local decision-making. Alongside a national decentralization program that re-established district-level councils, the government piloted a community-driven development project that went one step further by providing small grants to be administered by village development committees. This extension down to the village level aimed to establish more inclusive and accountable local decision-making infrastructure, rebuild trust, promote collective action, and provide minority groups (particularly women and youth) with experience in managing projects and making decisions within their community.

Household surveys, which covered participation in local decision-making, attitudes to minorities, and engagement in collective action, as well as demographic and socioeconomic information, were collected in late 2005 and again in mid-2009, along with village-level focus group discussions. In addition, three structured community activities (SCAs) were conducted in late 2009, shortly after GoBifo activities had ended, to capture any persistent impacts on collective action, participation of minorities, and elite capture. The SCAs were designed to measure how communities responded to concrete, real-world situations in three areas where GoBifo had sought to change behavior: (i) raising funds in response to a matching grant opportunity; (ii) making a community decision between two comparable alternatives; and (iii) allocating and managing an asset that was provided for free.

Institutional Change and Collective Action: There is no evidence that the program led to fundamental changes in local institutions or decision-making. Despite the fact that many women in treatment villages participated in GoBifo decisions, they were no more likely to voice an opinion in community meetings after the project ended or to play a leadership role in other areas. Similarly, the establishment of a democratically elected village development committee that carried out multiple projects did not lead treatment villages to be any more successful at raising funds in response to a later matching grant opportunity. Lastly, there were no program impacts on elite capture, although levels of capture were low in the research communities (at least as measured by the third SCA).

Source: Poverty Action Lab, <http://www.povertyactionlab.org/>

Some of the most revealing measurement results are one-time events conducted as experiments to test theories of corruption. These studies are based on carefully designed, highly detailed, and locally bound methodologies, and are rarely replicated or sustainable in the long term. Their purpose is to test theories about behavior, and to establish a strong correlation between intervention and outcome. They are helpful for informing the field of corruption measurement about what works in certain contexts, but they are not a useful vehicle for large-scale or replicable data collection.⁵⁰

⁵⁰See Annex 3 for more examples of policy experiments using objective measures of corruption.

Building Political Collusion: Evidence from Procurement Auctions (Coviello and Gagliarducci, 2010)

Researchers investigated the relationship between the time politicians stay in office and the functioning of public procurement. Data was collected on Italian municipal governments and all procurement auctions administered between 2000 and 2005. The primary finding is that one extra term in office deteriorates public spending. In fact, it decreases the number of bidders and, most importantly, the winning rebate. Interestingly, researchers also find that the probability that the same firm is awarded more auctions, or that the winning firm is local, increases with time in office. These results are compatible with the predictions of a model of favoritism in repeated procurement auctions, where time reveals collusive types, thus increasing the value of illegal connections at the expense of higher procurement costs.

Source: Institute for the Study of Labor, <http://www.iza.org/>

Theory-based Evaluation

In contrast to experimental impact evaluation, theory-based evaluation is based on the theory of change developed during M&E planning processes. The process of evaluation is built around the 'theory,' which is a set of assumptions about how an intervention achieves its goals and under what conditions.⁵¹ Evaluation methods assess the value and relevance of the project by "testing" the theory, i.e., exploring why and how projects cause results.⁵²

These kinds of evaluations are useful for complex, non-linear interventions with multiple factors at play, all of which contribute to final outcomes. They aim to support project improvements, build knowledge for generalizability and wider application, and support accountability.⁵³ A variety of qualitative and quantitative methodologies are used to establish the external outcomes of a project, many of which are also used for measurements of corruption, transparency, accountability, and integrity.

There are six key principles to a theory-based impact evaluation⁵⁴:

1. Map out the causal chain (theory of change)
2. Understand the context
3. Anticipate multiple outcomes
4. Evaluate impact using a credible counterfactual
5. Use facts as the basis for analysis
6. Use mixed methods (quantitative and qualitative)

Theory-based evaluation can be incorporated into experimental impact evaluations to deconstruct the "black box" of causality. However, it can also be used as the basis for non-experimental evaluations that aim to identify causal processes, and support contextual factors and causal mechanisms that influence outcomes. Rather than establishing cause and effect and estimating quantifiable impact, these non-experimental evaluations focus on making contributory claims about project activities and actual outcomes.

⁵¹Stern et al., *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID, 2012.

⁵²European Commission, *Evalsed Sourcebook: Methods and Techniques*, 2013.

⁵³UNDP, *Handbook On Planning, Monitoring And Evaluating For Development Results*, 2009.

⁵⁴Adapted from Howard White, *Theory-Based Impact Evaluation: Principles and Practice*, International Initiative for Impact Evaluation Working Paper, 2009.

Fictional Case: Making Contribution Claims about an Advocacy Campaign

An NGO provides funding to advocacy campaigns that aim to impact on budget expenditure. These campaigns do not work in isolation – a number of other contextual factors may play a role in bringing about an observed budget impact. Key impact evaluation questions are:

- Has the campaign plus these other contextual factors resulted in the observed budget impact?
- Was the campaign sufficient to bring about the budget impact?
- Was the combination of the other contextual factors without the campaign enough to bring about the budget impact?

If confirmed, the 'contribution claim' would be that the advocacy campaign 'worked' in that it made a difference and contributed to bringing about the observed impact. The Theory of Change for a campaign traces the logical sequence in the causal chain between the campaign's activities and the observed budget impacts, identifying other contextual factors that are needed for the links in this chain to work. The causal claim will then be two stages:

- First, the links in the causal chain must be shown to have happened and explained as to why they happened. This includes identification and discussion of contextual factors and other contributing causes that brought about each point in the sequence.
- Second, plausible rival explanations of why each link in the causal chain happened are identified and discussed in terms of their relative significance in bringing about impact.

Not all campaigns will work. Breakdowns in the expected sequence of impacts are also useful to help identify what could be done differently in the future.

Source: Stern et al., *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID, 2012.

The approaches to evaluation outlined below incorporate these principles to varying degrees in the process of assessing impact. The approaches are not so much focused on *measurement*, as they are focused on *understanding* the mechanisms of change, *capturing* unintended and negative outcomes, and *linking* outcomes back to project activities in a rigorous fashion. Findings from these types of analyses can contribute to the design of future projects in similar ways as the generalizable findings from impact evaluation.

Most Significant Change

The most significant change approach is a participatory means of evaluation in lieu of indicators.⁵⁵ It involves the collection of *change stories* from community members, practitioners, and organizational management working on an anti-corruption effort. These stories are meant to capture changes in behavior and/or policy, and function as initial drafts of impact statements. As stories are filtered up through levels of authority in an organization, the most significant of these stories are identified and discussed by panels of designated stakeholders or staff. The impact stories that are selected are then verified through standard data-collection methods such as site visits, interviews, and reviews of project documents.

⁵⁵Rick Davies and Jess Dart, *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*, 2005.

Outcome Mapping

Outcome mapping is an actor-centered approach that focuses relationships of actors, both within the project and in government, to track changes in behaviors that matter for outcomes. Outcomes are defined as changes in the behavior, relationships, activities, or actions of the people, groups, and organizations with whom a program works directly.⁵⁶ In terms of project-level evaluation, outcome mapping is an innovative, flexible approach to monitoring and evaluation of project-level outcomes. It stands in contrast to the linearity of the log frame approach, as the process of outcome mapping involves recognizing (unpredictable) opportunities, monitoring progress, and analyzing change throughout the life cycle of a project.

Outcome mapping recognizes the complexity of projects, incorporates participatory monitoring and evaluation practices, and aims to strengthen the adaptive capacities of stakeholders. However, the approach is labor-intensive, requires substantial capacity-building, and considers a much longer time-horizon than ordinary interventions.

Outcome Mapping in Transparency and Accountability

The AcT Programme in Tanzania is designed to increase government responsiveness and accountability through a strengthened civil society. The outcome mapping (OM) tool was used to shift the focus of civil society organizations from the output level (on training, workshops, carrying out pieces of research), to thinking about transformational change, which required a more nuanced contextual understanding, clear strategic thinking and calculated risk taking. Outcome mapping is an ideal approach to capturing impacts in a complex, multi-stakeholder environment where results are unlikely to be achieved in a linear fashion.

The challenge has been demonstrating change in a credible and consistent fashion within the constraints of a log-frame. Conventional indicators are good for clear, major steps, like national poverty monitoring systems, but don't often provide enough nuance about changes in attitudes of individual stakeholders, or the smaller initial steps (such as meaningful engagement with civil society, provision of information) that will contribute to the achievement of higher-level results (water points functioning, higher enrollment levels, etc.). The team realized mid-way through the project that they needed to articulate a theory of change (ToC) that would clarify what results they were aiming to achieve, and hence what kinds of indicators, qualitative and / or quantitative, would be appropriate.

The narrative reports written by partners (CSOs) every six months contain mini case studies, success stories, as well as most significant change stories. Partners observe changes, sense that they matter, and report them back as anecdotes in the context of longer narrative reports, but there might be a lack of certainty as to their value, as they can be seen as 'one off', with uncertain representativeness or replicability, nor do they fit into their conventional indicators. However, once they are collated alongside results from other partners, patterns of governance and accountability changes may appear (as well as outliers). And so the conventional indicators become 'containers' which can hold quantitative and qualitative information from any monitoring, including OM monitoring.

OM is not for the faint-hearted. It takes a lot of investment of time and energy. It requires teaching and learning for both donors and partners, and a one-size-fits-all model does not work. It generates a sizable amount of both qualitative and quantitative data that must be sorted through to find patterns and trends. Over time, however, it provides a detailed and systematic body of qualitative and quantitative evidence that takes us beyond anecdotes, and towards a nuanced understanding of what makes change happen.

Source: Kate Dyer, Making 'Evidence' the Plural of 'Anecdote': A 'Work in Progress' using Outcome Mapping and the Logframe in governance and accountability programming in Tanzania, 2012. www.outcomemapping.ca

⁵⁶Earl, S., Carden, F. & Smutylo, T. *Outcome Mapping. Building learning and reflection into development programs*, International Development Research Centre (IDRC), 2001.

Process Tracing

Process tracing is a qualitative research approach that investigates causal processes and mechanisms of projects, including anti-corruption interventions. It follows theory-based evaluation principles closely through a case study method. Evaluators work with stakeholders to identify intermediate and final outcomes, and then conduct a systematic assessment of whether project activities contributed to outcomes, which includes a “process induction” stage that seeks potential alternative causes. “Process verification” is then conducted to assess the extent to which claims of causality or contribution are supported by the evidence.⁵⁷

Anti-corruption Policies Revisited: Global Trends and European Responses to the Challenge of Corruption (ANTICORRP)

ANTICORRP is a large-scale research project funded by the European Commission’s Seventh Framework Programme. The project started in March 2012 and will last for five years. The research will be conducted by 21 research groups in sixteen countries. The project’s starting point is that, while the knowledge about the negative impact that corruption has on various aspects of human well-being (such as economic prosperity, health, life satisfaction, gender equality, social trust, poverty and political legitimacy) has been well established, knowledge about how corruption can be successfully fought by political means is much less developed. The fundamental purpose of ANTICORRP is to investigate and explain the factors that promote or hinder the development of effective anti-corruption policies and impartial government institutions. A central issue will be how policy responses can be tailored to deal effectively with various forms of corruption. Through this approach, advancing the knowledge on how corruption can be curbed in Europe and elsewhere is the primary aim. Special emphasis is laid on the agency of different state and non-state actors to contribute to the fight against corruption.

One of the project’s main objectives is to explain governance regime change as documented by time-series data, through global models developed through quantitative comparative analysis. The change process of countries in transition from one governance regime to another will be analyzed in-depth through qualitative comparative designs and (causal) process tracing that do not only focus on formal institutional development, but also on implementation and anti-corruption entrepreneurship. The case studies will trace the process and the mechanisms of change, the strategies of actors and the mechanism of altering the power distribution of particularistic societies leading to new equilibria.

Source : ANTICORRP (of European Research Centre for Anti-Corruption and State-Building), <http://www.againstcorruption.eu/anti-corruption-projects/anticorrp/>

⁵⁷Oxfam, *Process Tracing: Draft Protocol*, 2013

Other Evaluation Methods

Other evaluation options may incorporate elements of theory-based evaluation, or extend beyond the borders of causal inference to encompass costs and benefits, and multiple case studies. It is important to note that any of these design options may employ participatory approaches, and unlike traditional evaluations, they can be conducted both during *and* after implementation.

- **Contribution analysis** establishes a credible and plausible claim that a reform effort contributed to outcomes or impacts. It consists of six iterative steps to collect information, consider evidence, and suggest explanations: (1) Map out results chain based on project activities, (2) Develop theory of change, (3) Gather evidence and explore alternative explanations for results, (4) Assess the resulting contribution story, (5) Seek additional evidence, and (6) Revise the contribution story.⁵⁸
- **Cost-benefit analysis** attempts to determine the financial benefits associated with the outcomes or impacts of a reform effort. It can focus on individual or societal welfare, government performance, or other types of reform efforts.⁵⁹
- **Single- or multiple-case studies** involve the analysis of context (including political dynamics and enabling environments), causal processes, outcomes, and unintended results or unexpected consequences. Case studies use a variety of quantitative and qualitative methods to collect information, and as such, can provide a comprehensive understanding of what happened.⁶⁰

Another more extensive alternative to evaluation is to conduct comparative meta-analysis on selected groups of cases. Meta-analyses compare results from different studies in order to explain variation across countries, within countries, and the political dynamics that influence the enabling environment.

“That means reframing the question: how often does something work, by what criteria, and to what degree? Plus, who decides ‘what counts’ as working? If a transparency/accountability intervention only ‘works’ in some sense, let’s say, a third of the time – is that a success or a failure?”⁶¹

⁵⁸Julia Coffman, *UNICEF Advocacy Toolkit Companion*, 2010; Evalsed Sourcebook: Methods and Techniques, 2013.

⁵⁹Jesper Johnson, *Cost-effectiveness and cost-benefit analysis of governance and anti-corruption activities*, U4, 2014

⁶⁰Joachim Blatter and Markus Haverland, *Designing Case Studies: Explanatory Approaches in Small-N Research*, 2012.

⁶¹Jonathan Fox, *Think Piece: “Speaking Of International T/A Initiatives And CSO researcher Dialogue”*, November 2013.

Synthesis studies and systematic reviews evaluate all high quality research evidence relevant to one research question. They focus on accumulation and aggregation within a number of perspectives (statistical, theory based, ethnographic etc.), and may be used to build a generalizable theory about anti-corruption efforts.⁶²

Systematic Review of Anti-corruption Efforts⁶³

A systematic review was conducted on the effectiveness of micro-level anti-corruption strategies implemented in developing countries. The exclusion criteria were applied to nearly 6,300 papers and resulted in the inclusion of 14 studies in the synthesis of results. The findings suggested that effective monitoring programs have two important requirements: first, the programs must be implemented and monitored by a party desiring to lower corruption, and second, monitoring programs must be combined with either a financial or non-financial incentive program. The review presented seven practical and policy recommendations:

- Monitoring and incentives should be combined.
- The monitoring and incentives scheme must align with all involved parties' incentives and local-specific market structures.
- Community-level monitoring can be successful, but only when the community can punish corruption.
- Media can be a useful incentive for enforcing corruption reduction.
- Decentralization may be particularly successful where there are local capacity and high levels of participation.
- Decentralization is only successful when decision-makers and service providers are held accountable by program recipients.
- Non-governmental organizations can be useful tools in implementing programs that change the rules or alter monitoring and incentives schemes.

The most successful corruption-reduction strategies create a situation in which the potentially corruptible official chooses not to engage in corruption because the cost of corruption outweighs its benefits. This can be brought about by increasing both the probability of being caught and the punishment if caught. It can also be brought about by placing the corruptible decision in the hands of someone who faces a naturally higher cost of being corrupt.

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

⁶²Stern et. al, 2012.

⁶³Systematic reviews of qualitative research generally include four rigorous stages: a search strategy to locate studies; the application of an inclusion/exclusion criteria; assessment of effectiveness from the included studies; and synthesis of study findings.

Approaches to Measuring Effectiveness at the Project/Sector Level

Approaches at the project level employ a variety of techniques and tools to determine whether project goals are being met, and to capture and explain project outcomes. In particular, participatory monitoring and evaluation is a common technique for gathering data that is both relevant to the project and to stakeholders. It is important to remember that attribution is often not possible in complex interventions that involve advocacy, empowerment, and corruption risk. Most evaluations aim to explain contributions on short and medium-term outcomes.

Advocacy Campaigns

Advocacy campaigns attempt to influence policy by delivering evidence-based recommendations, and by encouraging decision-makers, stakeholders, and relevant audiences to support change. The focus of change efforts includes not just policy, but attitudes, institutional functions, power relations, and social inequities.

As with the measurement of corruption and its impacts, monitoring or assessing the effectiveness of advocacy efforts is not straightforward. The focus of assessment is important. Potential areas of assessment include raising the level of awareness, changing behaviors, reforming policies and law, or supporting coalition-building. The time frame for change might also extend well beyond the duration of the project. Context is dominant, in that it determines how, when, and why certain techniques will be used, and it means that dynamic contexts can result in fluctuating advocacy strategies and targets during implementation. Because of shifting priorities, traditional methods of evaluation are not always appropriate. The focus is on identifying potential contributions of advocacy efforts to outcomes, based on sound evidence and clear explanations of context. In some cases, participatory monitoring and evaluation and self-assessment provide valuable insight into change processes and policy outcomes.

Transparency International's Impact Monitoring Approach

In 2013, Transparency International developed and piloted a new approach to evaluating the impact of anti-corruption advocacy campaigns. The Impact Monitoring Approach is a twin track approach that consists of two complementary elements. The impact matrix is an analytical lens that helps to organise and structure the impact data which is captured by TI. It is organised around two main change areas which capture the main types of TI's impact:

- **POLICY CHANGE:** TI's contribution to policy changes in the public, private and civil society sector.
- **BEHAVIOURAL CHANGE:** TI's contribution to behavioural changes of individual people and institutions.

For each of the change areas above five levels were defined on the basis of the respective underpinning theories of change. That is, these levels reflect a progressive (albeit non-linear) trajectory toward higher impact levels, which is based on the assumptions of how change in these specific areas finally happens.

Impact reviews are deeper-dive impact assessments of TI projects/ initiatives conducted each year with the view to: (a) identify and analyse their areas and levels of impact and (b) capture lessons and promising approaches which can be replicated elsewhere.

Source: Interview with Rute Caldeira, Transparency International Secretariat, 2014.

The types of activities that can be monitored in advocacy efforts are varied, and depend on the nature of the specific project:

<i>Digital outreach</i>	<i>Voter education</i>	<i>Policy development</i>
<i>Media coverage</i>	<i>Briefings</i>	<i>Policymaker education</i>
<i>Media partnerships</i>	<i>Polling</i>	<i>Relationship building</i>
<i>Coalition building</i>	<i>Pilot projects</i>	<i>Litigation</i>
<i>Organizing rallies/marches</i>	<i>Policy analysis</i>	<i>Lobbying</i>

Evaluation of advocacy efforts tends to rely on non-experimental designs, as projects typically have non-linear results chains that shift and flow over time. Data-collection methods focus on tracking of media exposure and the public support of officials, but also rely on surveys, focus groups, interviews and other indicator/scorecard-driven assessments. The types of measurements that support evidence-based evaluations are listed in Table 6.

Table 14: Measurements of External Outcomes in Advocacy Efforts

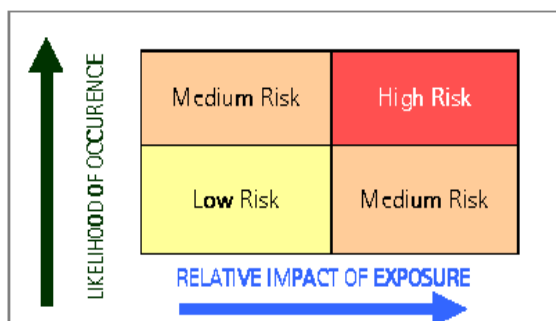
	Interim outcomes	Advocacy goals	Longer-term outcomes
What should be measured?	Organizational advocacy capacity Partnerships New advocates New champions Organizational or issue visibility Awareness Salience Attitudes or beliefs Public will Political will Constituency growth Media coverage Issue reframing Expert reference of materials	Policy development Placement on the policy agenda Policy adoption Policy blocking Policy implementation Policy M&E Policy maintenance New donors More or diversified funding	Improved services and systems Positive social and physical conditions Improved behavior of the stakeholders

Source: Adapted from UNICEF Advocacy Toolkit Companion, 2010

Mitigating Risks in Sectors

M&E approaches for anti-corruption interventions in sectors are primarily concerned with risk assessments and the mechanisms for monitoring both risks and corrupt practices. Risk is composed of the likelihood of a corrupt practice occurring and the subsequent impact of that corrupt practice. Assessments prioritize risks according to the characteristics and vulnerabilities of specific sectors such as education, health, agriculture, forestry, etc., and propose solutions to mitigate or eliminate them.

Figure 14: Risk Assessment as Combination of Likelihood and Impact



Source: Arthur G Blundell and Emily E Harwell, *Analyzing Corruption in the Forestry Sector*, Transparency International, 2010

Risk assessments focus on identifying the types of arrangements or practices that may lead to corrupt behavior, and may or may not include scores or ratings. They consist of evaluative data about organizational procedures, resource chains, and practices that can assist organizations in preventing corruption. Data is collected by a small team of experts through a variety of means, including interviews, surveys, observation, and then combined in the analysis stage with administrative data. Resulting data is not a measure of corruption, but rather, a measure of corruption risk. This data provides the basis for the corruption risk management action plan. The action plan identifies short-, medium-, and long-term priorities and indicators to manage the risks. The risk logs are the main M&E framework compliance and actions performed for preventing corruption.

Figure 15: Likelihood and Consequences of Risk Factors: Risk Cataloging

Risk Factor	Vulnerability	Likelihood of occurring	Economic consequences	Reputational consequences	Organizational consequences
Acute	Existing control measures do not offer any protection	Will most likely take place	Acutely high costs	Constant negative media attention. Criticism from donors and public.	Employees question the motives behind the policies of management and several resign
High	Control measures are weak and can be circumvented.	Is likely to happen at some point	High costs	Low confidence in the organization and management.	Significant internal discontent and reduction in staff morale
Medium	Control measures are good, but vulnerable to strategic attacks.	Might happen at some point.	Medium costs	Doubts about the organization, policies, and staff.	Reduction of staff morale in certain offices or departments
Low	Control measures provide adequate protection.	Little risk of occurrence	Low costs	Insignificant consequences for reputation.	Insignificant effect on morale or ethics.

Source: Global Advice Network APS, *Creating a catalogue for risk management. An interactive approach.*

Monitoring and evaluation mechanisms in sector approaches consist of a variety of tools, including anti-corruption instruments, data-collection methods, corruption/anti-corruption measurement methodologies, and financial tools prevalent within the sector. Much like with corruption measurement, selection of the most appropriate mechanisms depends on the purpose of the intervention and the available resources at hand.

Public Procurement Monitoring in Nigeria

Public and Private Development Centre (PPDC) of Nigeria designed and launched the procurement monitor's portal observatory as a central mechanism for collation, analysis and e-reporting of citizen-led procurement monitoring. The portal provides 24-hours access for virtual submission and collective analysis of procurement monitoring reports by registered monitors. It has a free online training tool, a blog for providing free legal advice to investigative journalists and monitors, and a virtual public procurement library. PPDC works with trained procurement monitors who are affiliated with NGOs or professional bodies.

Using a checklist, procurement monitors fill in their observations from procurement monitoring in the field and then submit the reports. PPDC collates the reports through the automated system on the portal and this provides statistical feedback that is submitted to the regulators, the Bureau of Public Procurement (BPP), as evidence of the level of compliance within Federal ministries and extra-ministerial departments with the public procurement process. Before the reports are submitted, they are turned into digestible content, linking procurement processes within ministries and also proffering solutions in areas where intervention is needed. This gives BPP a substantial starting point to enforce compliance of ministries with the reform process.

In addition to the procurement portal observatory, PPDC has a robust mailing list that is used to disseminate, share and receive information among stakeholders in the procurement process. The mailing list has frequently been used to mobilize citizens for action, for example, during the fuel subsidy removal in Nigeria. However, all these tools are enablers and depend on the activeness of citizens who are the main drivers of the procurement reforms.

Source: Claire Schouten, ICT for Open Contracting in Fragile and Conflict-Affected States, World Bank, 2012.

Interventions that are associated with particular corruption risk areas can be monitored and evaluated with a variety of anti-corruption mechanisms and instruments. Long-term outcomes in the health sector can be evaluated with commonly-used indicators of health and well-being. But intermediate outcomes associated with corruption and weak governance can be captured with any combination of instruments, tools, and methodologies. Indicators can be developed based on the issues areas associated with risk in the sector, and theory-based evaluation can be applied to determine the causal processes and mechanisms for change.

Table 15: Key Tools to Identify, Track and Measure Corruption Risks and Corruption in the Health Sector

Area	Issue	Tools to identify and track problems
General	Cross-cutting	<ul style="list-style-type: none"> Political economy analysis in the health sector Vulnerability to corruption assessments Value chain analysis Sector accountability assessment Value for money audits Analysis of governance in health care systems
Budget	Budget processes	<ul style="list-style-type: none"> Public Expenditure and Financial Accountability indicators (PEFA) Focus groups and interviews with public officials, recipient institutions, and civil society
	Payroll leakages	<ul style="list-style-type: none"> Public Expenditure Tracking Surveys and Reviews (PETS, PERS) Household surveys Focus groups with public officials and health workers
	In-kind leakages	<ul style="list-style-type: none"> Public Expenditure Tracking Surveys (PETS) Quantitative Service Delivery Surveys Facility surveys Focus groups with public officials, recipient institutions, and health workers
	Pharmaceuticals	<ul style="list-style-type: none"> WHO Good Governance in Medicines programme to assess transparency in drug supply and management International Drug Price Indicator Guide Internet based drug procurement databases
Individual Providers	Job purchasing	<ul style="list-style-type: none"> Official administrative records combined with facility surveys Interviews with public officials and former officials Governance and Anti-Corruption Country Diagnostic surveys
	Health worker absenteeism	<ul style="list-style-type: none"> Quantitative Service Delivery Surveys Surprise visits Direct observation Facility records Focus groups or interviews with facility heads and patients
Informal Payments	Informal payments	<ul style="list-style-type: none"> Household surveys (e.g. WB Living Standards Measurement Surveys and Demographic and Health Surveys (DHS)) Facility exit surveys and score cards Focus groups/interviews with providers/patients and health staff Governance and Anti-Corruption Country Diagnostic surveys
Corruption & Perceptions Experience	Perceptions of Corruption	<ul style="list-style-type: none"> World Bank Governance Indicators (Control of Corruption), TI Corruption Perception Index Governance & Anti-Corruption Country Diagnostic surveys (WB) National level perception surveys by CSO and others
	Experiences of corruption	<ul style="list-style-type: none"> AfroBarometer, LatinBarometer, EuroBarometer, TI Global Corruption Barometer National experience-based surveys Patient satisfaction surveys and report (score) cards Focus group surveys /studies

Source: DFID, *How To Note: Addressing Corruption in the Health Sector*, 2010. Adapted from M. Lewis and G. Pettersson, "Governance in Health Care Delivery: Raising Performance," *World Bank Policy Research Paper*, 2009.

In the case of the forestry sector, risks may be monitored at micro, macro, and organizational levels, by organizations located in- and outside the country, with datasets that span any number of countries. Indicators for evaluative purposes can be designed to capture changes in the nature or extent of the risk area, e.g., incidence of logging without community consent, amount of fees or taxes paid, amount of revenue reported, and incidences of falsified audits in one region. A baseline data-collection effort at the beginning of any anti-corruption project is essential to calculating changes in indicators. Care must also be taken to select appropriate data, given the type of risk and context in which it prevails.

Table 16: Anti-Corruption Monitoring Mechanisms for the Forestry Sector

Risk area	Anti-corruption instruments	Additional monitoring mechanisms
Regulatory		
Undue influence on forest laws and regulations; Forest zoning	MoF working groups for regulations (regs); Lobbying regs; Transparency regs for drafting of bills; Legislative ratification of bills/major regs; Well-advertised public comment periods; Freedom of expression and free press; Whistleblower protection; Ombudsman	National ethics board; NGO newsletters/reports on the legislative process for the bills/regs they follow; Annual checklists; Global Integrity Index; OECD; Government at a Glance; World Bank Governance Index
Licensing		
Preferential award of concessions and licenses	Procurement website; Government tender board/ procurement office; Accurate and unambiguous description of procurement and concession terms; Publication of bid proposal and decision criteria; Debarment for corrupt actors; Independent audits	TI CPI; Global Integrity Index; Local environmental NGO monitors occasional reports
Logging community land without consent	Grassroots engagement and awareness campaigns; Citizen complaint boards	WRI illegal logging indicators; Local community- based organizations occasional reports
Timber supply		
Illegal logging	Chain of custody timber tracking; Independent observer at timber checkpoints; GIS monitoring; Independent field monitoring; Citizen complaint channels; Industry codes of conduct; Incentives for MoF employees	WRI; Chatham House illegal logging indicators; Mirror statistics for production/trade
Illegal use of labor	Citizen complaint channels; Labor review boards	International Labor Organization (ILO) Gaps in Workers' Rights; Local labor-NGO reports
Illegal use of (unaccountable/ armed) security forces	Citizen complaint channels; Voluntary private sector agreements on use of security; NGO and grassroots field observations; engagements with local communities	Freedom in the World; Global Integrity Report; Human rights NGOs' occasional reports
Officials use government resources for their own logging companies	Public access to annual audit of uses of government resources; Citizen complaint channels	Public Expenditure and Financial Accountability (PEFA) Assessments; National auditing body reports
Log transport without proper documents	Chain of custody; Independent observer at checkpoints	WRI; Chatham House; Local environmental NGOs' occasional reports
Use of illegal wood in processing industry	Chain of custody; Independent observer at entry points	WRI; Chatham House; Local environmental NGOs' occasional reports
Smuggling	Chain of custody; NGO undercover investigations; Wood balance analysis; FLEGT or similar import requirements	WRI; Chatham House; Local environmental NGOs' occasional reports

Table 16: Anti-Corruption Monitoring Mechanisms for the Forestry Sector (continued)

Reporting		
Transfer pricing	Mirror statistics; Customs reporting reforms; Training of customs agents to recognize high-value species.	FAOSTAT; ITTO Market Information System
Under-reported volume or value (domestic tax evasion)	Chain of custody	FAOSTAT; ITTO Market Information System; Chatham House; WRI
Laundering illegally sourced wood into the legal supply chain	Transparent annual reporting by wood industry	WRI; Chatham House
Failure to fully and accurately report revenues; Excessive credits for fees and taxes	Transparent online payment systems at MoF	Global Integrity Index; PEFA; National auditing body reports; (For publicly traded companies) company internal audits
Failure to satisfy financial obligations to communities	Transparent reporting of payments; Annual audits of community development funds; Citizen complaint channels	Grassroots advocacy groups; Citizen whistleblowers
Revenue		
Non-payment of fees	Transparent online payment systems; Annual audits of MoF accounts	Global Integrity Index; Open Budget Index; World Bank Governance Index; PEFA; National auditing body reports; company internal audits
Use of sweep accounts to make overnight loans using deposits of forestry fees	Requirement that fees be paid directly to Ministry of Finance/national bank; Transparent online payment systems at MoF and Ministry of Finance	PEFA; National auditing body reports
Lack of oversight; Sanction for unpaid taxes; Late transfers of forest revenues	Transparent online payment systems at MoF; Annual audits of MoF accounts	PEFA; National auditing body reports
Falsified audits; Failure to report irregularities to proper authorities	Public access to audits	PEFA; National auditing body reports; (For publicly traded companies) company internal audits
Neglect of Know-Your-Customer due diligence/Suspicious Transactions Reports (STRs)	Summary reporting of STRs from each institution (publicly available)	OECD; Financial Action Task Force (FATF); National financial intelligence body
Money laundering of proceeds from illegal logging to support political campaigns	Assets disclosure regulations and reporting; Campaign financing reporting (publicly available)	OECD; FATF; Freedom in the World; Global Integrity Report; National financial intelligence body; National auditing body; National election oversight body; Candidate wealth reporting body

Source: Arthur G Blundell and Emily E Harwell, *Analyzing Corruption in the Forestry Sector*, *Transparency International*, 2010

Empowerment and Capacity Building

Empowerment is a difficult concept to define because it involves a complex mixture of values, knowledge, behavior, and relationships.⁶⁴ It might be based on global rights that are implemented in very diverse legal and societal contexts. It may be associated with the capacity to make choices and transform choices into desired outcomes.⁶⁵ It might be measured using proxy indicators such as resources (the conditions under which choices are made); agency (the process by which choices are made); and achievements (the outcomes of choices).⁶⁶ It is often subsumed into capacity-development projects that aim to facilitate active participation and monitoring of service delivery and policy-making. These activities are buttressed by knowledge, access, and political agency.

⁶⁴Governance and Social Development Resource Centre, <http://www.gsdrc.org/go/voice-and-accountability/>

⁶⁵Ruth Alsop and Nina Heinsohn., *Measuring Empowerment in Practice: Structuring Analysis and Framing Indicators*, World Bank, 2005.

⁶⁶Naila Kabeer, *Reversed Realities: Gender Hierarchies in Development Thought*, 1994.

Monitoring empowerment involves tracking changes in relationships. A “fit for purpose” set of empowerment indicators is one which provides sufficient description of changes in power relations to frame and prompt in-depth analysis of those changes in ways that will lead to improved empowerment interventions and help hold decision makers accountable for the impacts they have on people’s lives.

Source: Jeremy Holland and Laurent Ruedin, *Monitoring and evaluating empowerment processes*, Swiss Agency for Development and Co-operation, 2012

Power dynamics play an important role in the success of initiatives that aim to empower marginalized groups. Traditionally excluded groups do not benefit from inclusion in decision-making processes unless their material conditions are improved as a result of participation. Similarly, changing the flow and amount of resources may stimulate opposition from formerly powerful actors, and serve to hamper citizen efforts to effect change. Participatory mechanisms for empowerment must take care not to place participants in disadvantaged positions that threaten their already fragile means of survival.

Thai Youth Anti-Corruption Network

Sponsored by United Nations Development Programme (UNDP) in Thailand, the Thai Youth Anti-Corruption Network is an active group of more than 3,500 students from over 90 universities across Thailand on a mission to eliminate corruption in Thai society through the empowerment of young people. The campaign was built to get students to take a personal pledge to “refuse to be corrupt.” The campaign was created by students and implemented by students, and the results are the thoughts and ideas of those students.

The strategy of the network involves several different activities:

- Raise youth awareness about corruption and its negative impacts.
- Form an inter-university student network (via Facebook) as a platform for active participation and engagement in promoting integrity and preventing corruption.
- Train students on the use of social media as an advocacy tool and teach them interpersonal skills needed to share key messages with their peers.
- Establish a national brand, a logo, and key messaging that will become nationally recognizable.
- Partner with Thai universities to further expand the network and plan campus-wide/inter-university activities to promote integrity and prevent corruption.

This is a one-of-a-kind initiative because it is completely ‘bottom-up’ — the direction of the program is designed entirely by students. The campaign gave them the tools and capacity needed to build their own network.

The anti-corruption campaign has continued to attract attention from other institutions and networks, most notably Thailand’s private sector Anti-Corruption Network (ACN). On International Anti-Corruption Day 2,000 university students came to Bangkok — pouring out of mass transit stations dressed in trademark “Refuse to be Corrupt” blue t-shirts. They came from all over Thailand. More than 500 students came from the Southern provinces of Thailand. 23 universities participated in the anti-corruption themed art exhibition at the Bangkok Art and Culture Center, in partnership with the National Anti-Corruption Commission (NACC).

Capacity is defined as “the ability of individuals, organizations and societies to perform functions, solve problems, and set and achieve objectives in a sustainable manner.”⁶⁷ A results-based approach to assessing capacity building includes three levels of measurement:⁶⁸

1. Impact: Change in people’s well-being
2. Outcome: Change in institutional performance, stability and adaptability
3. Output: Activity completed or service provided based on capacity development core issues (institutional arrangements, leadership, knowledge, and accountability)

Both public sector institutions and communities can engage in capacity building activities, with similar outcomes. However, the types of outputs differ depending on the activity roles of individuals in anti-corruption efforts.

Approaches to Measuring Effectiveness at the Institutional/Organizational Level

Because monitoring and evaluation at the institutional level is concerned with organizations, there is a clear focus on processes and capacities. In anti-corruption authorities, measurement is concerned with outputs (e.g. increased capacity) and performance, but less so on outcomes and impacts because of the difficulty of attribution. Institutional integrity applies to all kinds of organizations and agencies in the public sector. Its focus is organizational risk, actual levels of corruption, and the ethics or behavior of public officials.

Anti-corruption Bodies

The monitoring and evaluation approaches employed by anti-corruption bodies should be reflective of their organizational mandates, which ordinarily consist of a combination of prevention, investigation, and/or prosecution. Monitoring consists of formal process of monitoring performance against objectives and aims in these areas. Evaluation is a less distinct area, as the factors that contribute to outcomes relevant for anti-corruption bodies are distributed across several different entities, including law enforcement, judiciary, supreme audit institutions, financial investigative bodies, and others. Many of the complaints about the ineffectiveness of anti-corruption bodies stem from the inappropriate expectations of the efficacy of these agencies on levels of corruption. It is highly unlikely that this is possible through the actions of just one government body.

At a practical level, in order to strengthen the capacity of anti-corruption agencies to more effectively discharge their mandates, it is necessary to first assess the existing capacity, which consists of three levels:

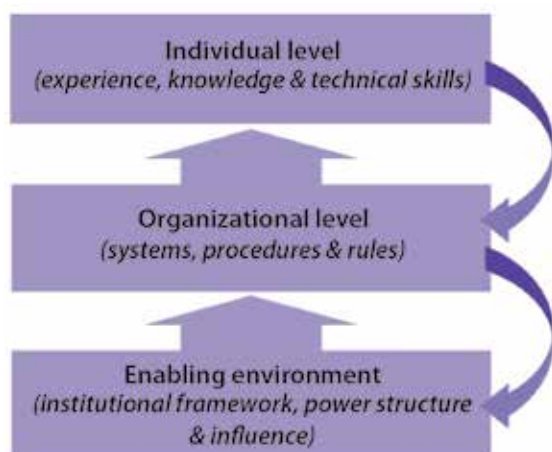
- the enabling environment (social, economic and political context including political will, institutional arrangement and coordination mechanisms, legal framework and the clarity of mandates, the level of independence, availability of human and financial resources, and oversight over the ACA);
- the organizational level (e.g., vision, policies, procedures, leadership, planning, management, monitoring and evaluation and other business processes that are essential for effective performance of the agencies);
- the individual level (e.g., staff skills).⁶⁹

⁶⁷UNDP, *Capacity Development Practice Note*, 2007.

⁶⁸UNDP, *Measuring Capacity*, 2010.

⁶⁹UNDP, *Practitioners’ Guide: Capacity Assessment of Anti-corruption Agencies*. New York: UNDP, 2011.

Figure 16: Capacity Assessment of Anti-corruption Agencies (UNDP Methodology)



The other important aspect of evaluating institution at the organization level is performance monitoring and outcome evaluation.

Table 17: Selected Indicators for Performance Monitoring and Outcome Evaluation of Anti-corruption Bodies

Outputs of ACA	Complaints: Number of complaints leading to preliminary or full investigation annually, average time to respond to a complaint
	Investigations: Number of investigations completed annually, number resulting in disciplinary action, number referred for prosecution, ratio of completed cases to total cases, average length of investigation
	Prosecutions: Number of prosecutions and of referrals dropped without prosecution annually, average trial time, conviction rate, average value of assets recovered and/or of gratification involved, percentage of corruption cases as percentage of all criminal cases, corruption case conviction rate compared to criminal case conviction rate
	Education function: Number of trainings, outreach events, and awareness campaigns.
	Systems Review Functions: number of reviews conducted, number of recommendations made
Outcomes	Education function: percentage of increase in understanding from pre-and post-training knowledge surveys, percentage of positive responses to trainings, percentage of awareness of respondents after awareness campaigns, percentage of trust/faith in ACA, number of requests from private firms for free corruption prevention advice
	Prosecutions: Number of persons prosecuted (corruption and related offenses), Number and rate of convictions

Source: Francesca Recanatini, *World Bank*, 2014

M&E systems can certainly improve performance and increase accountability, but they can also protect anti-corruption bodies from political pressures and interference.⁷⁰ M&E data can be used to manage expectations by painting a realistic picture of what can be accomplished, and when data is released to the general public, it can establish clear downward accountability to citizens and civil society.

Table 18: Country Examples of Indicators from Anti-corruption Bodies⁷¹

Mongolia	Number of cases investigated and solved under the law, percentage of public officials who submitted income and asset disclosure forms, number of actions plans adopted by public organizations and local governments
Brazil	Number of internal investigations completed and penalties enforced, number of companies suspended or debarred, number of web data portals and number of visits, number of training delivered and of participants
Indonesia	Number of cases investigated, conviction rate in corruption prosecutions

Source: Francesca Recanatini, *World Bank*, 2014

Because weak capacity is a core issue facing many anti-corruption bodies, capacity assessment exercises are an integral part of monitoring and evaluation exercises. But it should be noted that capacity assessment is part of a long-term process of capacity building. Similar to other forms of measurement, capacity assessment feeds into further action. The focus for anti-corruption bodies is function or core capacities that are needed for daily management of the agency, and technical capacities associated with professional anti-corruption expertise, such as forensic accounting skills, law/legal expertise, surveillance best practices, knowledge of finance and procurement systems and vulnerabilities, etc.⁷²

Institutional Integrity

Institutional integrity assessments are applied at the agency or organizational level, and are often conducted by specialized anti-corruption bodies. These assessments are similar to risk assessments in that they aim to identify vulnerabilities to corruption and actual levels of corruption. But they go further to examine organizational ethics and administrative culture, which involves study of the values and behaviors of public officials, and the constraining rules that attempt to mitigate risks or conflicts of interest.

⁷⁰Francesca Recanatini, Senior Economist, World Bank.

⁷¹For additional suggested indicators on the outputs, outcomes, and impact of ACAs, please see: Johnson et al., *How to monitor and evaluate anti-corruption agencies: Guidelines for agencies, donors, and evaluators*, Chr. Michelsen Institute (CMI), 2011

⁷²UNDP, *Practitioners' Guide: Capacity Assessment of Anti-Corruption Agencies*, 2011.

Table 19: Country Examples of Institutional-level Assessment Tools

Slovenia	Integrity plans in Slovenia employ both quantitative and qualitative methodology to estimate the level of corruption risk in an institutional environment and to propose solutions to alleviate that risk. Each public entity must create a project group of 5-7 individuals who will design and implement the integrity plan, which must consist of the following information: (1) Assessment of corruption exposure of the institution, (2) Personal names and work posts of individuals responsible for the integrity plan, (3) Description of typical work processes and decision-making method, including a corruption risk exposure assessment and proposed improvements regarding integrity, and (4) Measures for timely detection, prevention, and elimination of corruption risks. ⁷³
Moldova	The objectives of Corruption Risk Assessment in Moldova are to identify the institutional factors that might facilitate corruption, and to draw up recommendations for elimination or mitigation. The assessment is conducted in three stages: (1) assessment of legal and structural preconditions, (2) assessment of corruption risks, and (3) drafting of an integrity plan. ⁷⁴
Philippines	Integrity Development Review (IDR) in the Philippines is a preventative measure against corruption under the responsibility of the Office of the Ombudsman. It consists of a systematic assessment of the agency's corruption-resistance mechanisms and its vulnerabilities to corruption. It proceeds in several stages: (1) assessment by officials of agency vulnerabilities, (2) survey of officials' knowledge and experience with integrity measures, (3) evidence-gathering to support assessments. The IDR is followed by the Corruption Vulnerability Assessment Process, which draws up an integrity plan to combat the vulnerabilities. ⁷⁵
South Korea	The Integrity Assessment (IA) measures the levels of corruption and corruption risks in the public sector, through a survey of ordinary citizens and public officials who use public services that are considered prone to corruption. The assessment criteria consist of various factors including actual experiences with corruption and bribery, adequacy of regulation, information disclosure, and organizational culture. The components of IA consist of external integrity of service users and internal integrity assessments, in which public organization employees evaluate the integrity of their own organizations. ⁷⁶

⁷³Stephanie E. Trapnell, *Commission on the Prevention of Corruption Slovenia: A Review of the Effectiveness of Anti-Corruption Agencies*, World Bank, 2010.

⁷⁴Maira Martini, *Overview of Integrity Assessment Tools*, Transparency International and CHR Michelsen Institute, 2012.

⁷⁵Martini 2012.

⁷⁶Arsema Tamyalaw, *A Review of the Effectiveness of the Anti-corruption and Civil Rights Commission of the Republic of Korea*, World Bank, 2010.

Integrity Pacts

The Integrity Pact (IP) is a powerful tool developed by Transparency International (TI) to help governments, businesses and civil society fight corruption in public contracting. It consists of a process that includes an agreement between a government or government agency ('the authority') and all bidders for a public-sector contract, setting out rights and obligations to the effect that neither side will pay, offer, demand or accept bribes; nor will bidders collude with competitors to obtain the contract, or bribe representatives of the authority while carrying it out. An independent monitor who oversees IP implementation and ensures all parties uphold their commitments under the pact brings transparency and invaluable oversight to all stakeholders in a contracting process, from the authority to the general public.

An assessment process was conducted in India in 2011, in which various stakeholders were consulted through personal interviews, telephone, e-mails and by post. Attempts were made to identify financial gains and success stories supported by case studies from various public sector undertakings (PSUs), as well as to suggest recommendations for improved implementation.

It was found that almost all IP compliant PSUs, Independent External Monitors (IEMs) and a substantial number of bidders feel that IP has helped in making procurement process more transparent but there have also been non-responsive cases. It was also found that the general awareness of IP among bidders is low and the IP compliant PSUs are the ones that need to share a blame for this. Hence, there is a need to widen the level of awareness among bidders by organizing more focus group discussions, workshops etc. In terms of financial impacts, gains from re-tendering under IPs were substantial.

Source: Transparency International, Integrity Pacts in the Water Sector: An implementation Guide for Government Officials, 2013; Ashutosh Kumar Mishra and Nakul Gupta, Assessment of Integrity Pact (IP) In IP Compliant Public Sector Undertakings, Transparency International India, 2012.

Approaches to Measuring Effectiveness at the National Level

National-level monitoring and evaluation spans a plethora of processes, organizations, and expectations, and as such, is often performed over longer periods of time by larger teams of individuals. Self-assessment is a common approach, given the high-profile nature of evaluations and potential for politicization of results. However, there are also cases of evaluations performed jointly by two or more countries, facilitating the possibility of capacity building, knowledge transfer, and validation of results.

National Anti-corruption Strategies/Policies

National anti-corruption strategies are comprehensive plans for governments to tackle corruption and the issues driving corrupt practices. Robust strategies would include M&E indicators that were developed in the planning process. Monitoring and evaluation of these strategies occurs indirectly through National Integrity Systems (NIS) Assessments by Transparency International, and Public Sector Integrity Assessments by the OECD. These assessments are driven by external donors, but conducted jointly with country partners, and take a whole-systems approach to anti-corruption mechanisms within government.

OECD Public Sector Integrity Assessments

The OECD helps countries review and modernize their integrity framework by mapping out good practices and developing principles, guidelines and tools. The approach focuses on mapping “at risk” areas vulnerable to misconduct, fraud and corruption. The Integrity Framework is a comprehensive approach to view integrity management within government. It combines:

- The integrity management systems: the instruments, processes and actors within public sector organizations to stimulate and enforce integrity and prevent corruption.
- The integrity context (or supporting environment) that can have an impact upon the integrity of the members in public sector organizations.

The Public Sector Integrity Review process is linked to the corruption prevention chapter of the UNCAC. It is a systematic assessment by practitioners, combined with peer learning from other countries, which generates proposals for action building on international good practice, based over a decade of experience of the OECD in this field. It is considered direct support to a government’s anti-corruption agenda.

A recent 2011 review process in Tunisia assessed the legal and institutional frameworks against corruption, including the coordination of controls, specific corruption prevention measures, and vulnerabilities in public procurement. The results of the assessment were a map of main risk areas for corruption, identification of counter-measures and the sequencing of their implementation over time (to avoid a wish list), and the involvement of main stakeholders to facilitate implementation.

Source: Elodie Beth, Strengthening the corruption prevention framework, OECD, 2012.

By contrast, direct M&E approaches of national anti-corruption strategies are rarely conducted, even by agencies specializing in anti-corruption oversight. Annual reports by these agencies can be vehicles for reporting on midterm and annual progress, but data-collection and analysis are inconsistent. Data often cannot be collected through the administrative systems already utilized by government agencies, but in cases where oversight bodies have sufficient resources, qualitative methodologies may be employed to validate results. Rather than country-level monitoring of anti-corruption strategies, a more common approach is institutional risk assessment and corruption monitoring within specific sectors.

Kuala Lumpur Statement on Anti-Corruption Strategies

To counter the general failure to monitor impact of national anti-corruption strategies, high level representatives of anti-corruption authorities as well as national planning authorities from the South, East and Southeast Asia and anticorruption experts from around the world gathered in Kuala Lumpur in October 2013 to discuss a set of guidelines that could instruct the process of developing, designing and implementing sustainable anti-corruption strategies.

A key meeting outcome was the agreement on a set of guidelines or indicators in three areas of design and implementation:

- Anti-Corruption Strategy Development Process
- Anti-Corruption Design & Content
- Anti-Corruption Strategy Monitoring & Evaluation

Source: <http://www.anti-corruption.org/index.php/en/component/content/article/82-news/218-un-conference-produces-anti-corruption-strategy-recommendations-for-asia>

International and Regional Conventions/Treaties

Convention monitoring tools are mechanisms designed to monitor the extent to which national governments fulfill their commitments under international conventions and action plans.⁷⁷ They may be legally binding or voluntary, and they are conducted through a variety of methodologies, including self-assessments, expert reviews, peer reviews, on-site visits, civil society monitoring, and the publication of recommendations for improvement. Compliance monitoring is intended to be a pressure mechanism on national governments to meet their obligations and ultimately, strengthen local anti-corruption systems through enactment of robust legislation.

In addition, assessments can also serve to:

- Highlight difficulties encountered by states in implementing the standards and help identify specific needs for technical assistance.
- Promote international cooperation in the fight against corruption.
- Promote information exchange on successes, good practices and experiences in applying the standards.

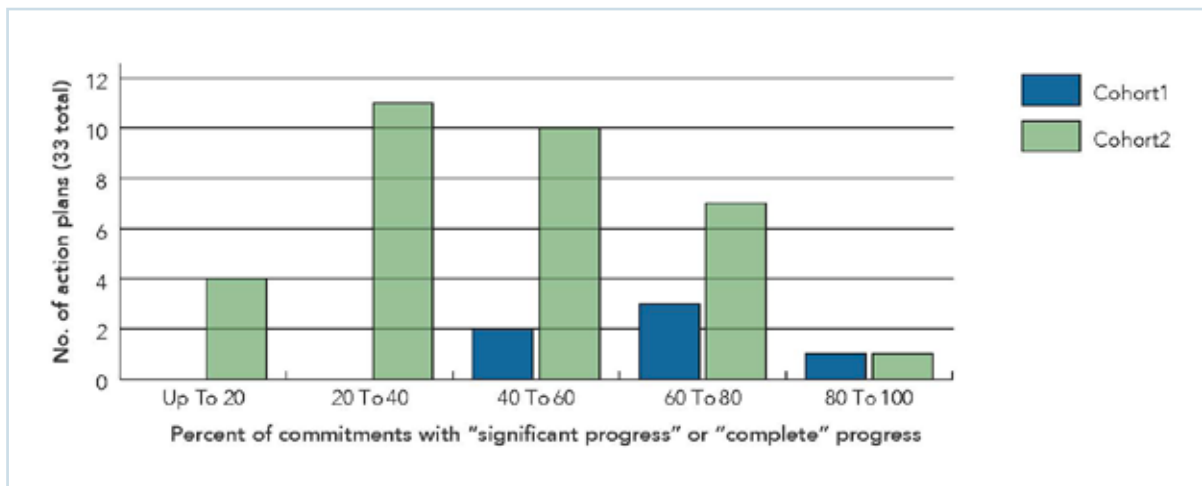
Source: Andy McDevitt. *Monitoring Conventions Topic Guide*, Transparency International, 2012.

⁷⁷Andy McDevitt. *Monitoring Conventions Topic Guide*, Transparency International, 2012.

Open Government Partnership Independent Review Mechanism

The Open Government Partnership is a new global, multi-stakeholder effort to make governments better. OGP aims to secure concrete commitments from governments to drive open-government reform and innovation at the country level, in an effort to stretch countries beyond their current baseline in the areas of transparency, accountability, and citizen engagement. The Independent Reporting Mechanism (IRM) is a key means by which all stakeholders can track OGP progress in participating countries. The IRM produces biannual independent progress reports for each country participating in OGP. The progress reports assess governments on the development and implementation of OGP action plans, tracks progress in fulfilling open-government principles, and make technical recommendations for improvements.⁷⁸

Figure 10: OGP Commitments with Significant or Complete Progress⁷⁹



The United Nations Convention Against Corruption (UNCAC) Review Mechanism

The UNCAC provides a set of standards, measures and rules that all countries can apply in order to strengthen their legal and regulatory regimes to fight corruption. It calls for preventive measures and the criminalization of the most prevalent forms of corruption in both public and private sectors. The UNCAC Review Mechanism is designed to assess the extent to which signatory parties comply with the provisions of the convention. The review process combines a self-assessment with a governmental peer-review country being reviewed by the governmental representatives of 2 other countries. On-site visits and civil society participation are at the discretion of state parties.⁸⁰ As of September 2014, there are a variety of country assessments available online: Self-assessment checklists (13), Executive Summaries (70), and Country Reports (30).

⁷⁸Open Government Partnership Independent Review Mechanism: <http://www.opengovpartnership.org/independent-reporting-mechanism>

⁷⁹Executive Summary of OGP IRM Technical Paper 1: http://www.opengovpartnership.org/sites/default/files/attachments/Technical%20paper%201_Executive%20summary_final.pdf

⁸⁰Comprehensive Self-Assessment Checklist on the Implementation of the United Nations Convention Against Corruption: <https://www.unodc.org/unodc/en/treaties/CAC/self-assessment.html>

Bangladesh and the UNCAC Gap Analysis

The UNCAC gap analysis sheds light only on parts of the reform agenda needed to improve governance in Bangladesh. It has to be seen as one piece of the puzzle. The sequence of steps taken by the caretaker government reveals the strategic approach behind UNCAC implementation: the first gap analysis was followed by a needs assessment to prepare a strategy for further UNCAC implementation. This process called for (a) updating and disseminating the first analysis, (ii) building the capacity of relevant public officials, (iii) enhancing Bangladesh's activities in the UNCAC working groups at the international level, and (d) putting more emphasis on mutual legal assistance and coordination of stakeholders. These requirements were partly addressed by putting in place the following:

- the second gap analysis,
- further legislative changes,
- capacity-building trainings,
- a national public procurement project,
- an action plan for UNCAC compliance, and
- the development of a National Integrity Strategy.

All of these steps are laudable and show the government's keen interest in addressing corruption in a holistic way. However, the emphasis on training, legislation, and (rather technical) action plans could be seen as an indicator that the anti-corruption drive relies on the usual technical fixes. It is important that there is now a paper trail of reform obligations which have been informed in part by UNCAC. This may help sustain the reforms across government cycles. However, this is only likely if key actors in government and state institutions internalize the commitments made under the Convention and are constantly reminded of them by outside stakeholders. UNCAC could thus be used as a tool for facilitating dialogue between the government and citizens about better governance. The value of UNCAC lies in the reinforcement it provides for the reform process than in the specific content it proposes, which is only partly consistent with national reform needs. It is important to take repeated reality checks to confirm that technical assistance needs identified under UNCAC respond to reform priorities.

Source: Hechler et. al, Can UNCAC address grand corruption? A political economy analysis of the UN Convention against Corruption and its implementation in three countries, Chr. Michelsen Institute (CMI), 2011.

Monitoring the OECD Convention on Combating Bribery of Foreign Public Officials in International Business Transactions

The OECD Anti-bribery Convention establishes legally binding standards to criminalize bribery of foreign public officials in international business transactions. The monitoring process involves self-assessment, on-site visits, peer reviews, plenary discussions, and the publication of country performance reports.⁸¹

Elements of Phase 3 Evaluation (2014)

- Appointment of two countries to act as lead examiners.
- Replies to an evaluation questionnaire by the country being evaluated.
- On-site visit to the country being evaluated.
- Preparation of a preliminary report on country performance.
- Evaluation in the Working Group on Bribery.
- Adoption by the Working Group of a report, including recommendations, on country performance.

Council of Europe Group of States against Corruption (GRECO) Monitoring

Council of Europe has developed a number of legal instruments dealing with matters such as the criminalization of corruption in the public and private sectors, liability and compensation for damage caused by corruption, conduct of public officials and the financing of political parties. GRECO is responsible for monitoring compliance with these standards through a process of mutual evaluation through on-site country visits and drafting of evaluation reports with specific recommendations. Measures taken to implement recommendations are subsequently assessed by GRECO under a separate compliance procedure.⁸²

Table 20: Number of Countries (40 Total) with Specific Themes Emerging from Recommendations and Observations Generated from GRECO Reporting Mechanisms

	Recommendations	Observations
Anti-corruption Strategy	9	1
Prevention/ Risk evaluation	10	2
Transparency	13	0
Control including by the Ombudsman	7	4
Statutory Rules/ Code of Ethics	16	1
Recruitment	5	0
Evaluation/career	5	0
Training/Awareness	20	0
Conflicts of interest, Incompatibilities – Accessory activities	12	2
Declaration of assets and interests	4	0
Rotation	8	3
Pantouflage (Post-employment practices)	14	0
Gifts	9	0
Reporting/Protection of whistleblowers	18	0
Disciplinary proceedings	3	1

Source: Author, using data from Group of States against Corruption (GRECO): *Lessons learnt from the three Evaluation Rounds (2000-2010)*, 2012.

⁸¹ Country Monitoring of the OECD Anti-Bribery Convention: <http://www.oecd.org/corruption/countrymonitoringoftheoecdanti-briberyconvention.htm>

⁸² GRECO Evaluations: http://www.coe.int/t/dghl/monitoring/greco/evaluations/index_en.asp

Mechanism for Follow-Up on the Implementation of the Inter-American Convention against Corruption (MESICIC)

The Organisations of American States (OAS) Convention obliges States Parties to implement a series of measures to prevent, detect, prosecute, and eradicate corruption in the public sector as well as to promote, facilitate, and regulate cooperation between State Parties on these matters. The monitoring process involves self-assessments and civil society input. Expert review sub-groups conduct a review with the State's response, involving meeting with government and civil society.⁸³

Monitoring of the ADB/OECD Anti-corruption Plan for Asia Pacific

The Action Plan sets out a series of goals and standards for sustainable safeguards against corruption in the economic, political and social spheres of the countries in the region. Reforms under this mechanism are conducted in implementation cycles:

1. At the beginning of each implementation cycle, each country identifies up to three priorities for reform and develops specific reform projects to tackle the identified weaknesses.
2. The implementing projects are discussed during Steering Group meetings to exchange experience with partners that have implemented similar reforms in the past or conduct reforms in linked areas.
3. At the end of each implementation cycle, the countries present the achievements and difficulties that were encountered in the implementation of the projects to provide feedback to the members of the Initiative and to share the experience gained in the implementation of the projects.⁸⁴

The African Peer Review Mechanism (APRM)

The APRM is a mutually agreed program, voluntarily adopted by the member states of the African Union, to promote and re-enforce high standards of governance in the areas of: democracy and political governance, economic governance, corporate governance and socioeconomic development. The review process includes country self-assessments, onsite visits by expert review teams who consult with government, private sector and civil society representatives, and the development of country reports and action plans.⁸⁵

There are four types of review:

- A base review, which is the first country review carried out within 18 months after a country becomes a member of the APRM.
- A periodic review that takes place every two to four years.
- A member country may, for its own reasons, request a review outside the framework of the periodically mandated reviews.
- Early signs of impending political and economic crisis in a member country could also be sufficient cause for commissioning a review.

⁸³ Mechanism for Follow-Up on the Implementation of the Inter-American Convention against Corruption: http://www.oas.org/juridico/english/mesicic_intro_en.htm

⁸⁴ Implementation of the Anti-Corruption Action Plan: <http://www.oecd.org/site/adboecdanti-corruptioninitiative/implementationoftheanti-corruptionactionplan.htm>

⁸⁵ African Peer Review Mechanism (APRM): <http://aprm-au.org/>

Lessons Learned from the Trenches (Practitioner Voices)

This chapter presents lessons learned by practitioners in the field of corruption and anti-corruption measurement. It addresses the need for complementarity in approaches, types of data, and levels of analysis. It also includes discussion of purpose, context, and credibility of measurement initiatives.

The Governance Data Alliance, a group of NGOs, governments, firms, and donors working to improve the quality, availability, breadth, and use of governance data, describes the current status of affairs as vastly insufficient when it comes to the production and usage of high-quality governance data: “Producers rarely know who uses their data; users have no way of signaling to producers what they want and need; and donors have no idea what the return on their investments is.” Box 1 below captures some of the major lessons learned by several practitioners. The lessons learned in this chapter are based on interviews with practitioners working in the fields of corruption, transparency, accountability, and integrity in countries all over the world. Over 25 individuals graciously provided their valuable insights during semi-structured interviews over audio calls. They included academics, researchers, policymakers, donors, and in-the-field experts. Topics ranged from the quality and relevance of corruption measurement, to the various methods for collecting and analyzing data, and the importance of capturing impact and using the data to establish accountability. Interviewee comments are captured in italicized text and quotation marks throughout the chapter.

Data Should Not Be Produced in a Vacuum

For much of the last two decades, the production and use of data on corruption and anti-corruption have been dominated by large donor organizations and isolated academic researchers. Data has been generated to assess progress, explore ideas, determine aid allocations, and evaluate effectiveness, all of which might take place without the involvement of local stakeholders. But data is a kind of information that can wield enormous power over those being assessed. Practitioners agreed that data must be shared to check that power. “Enough of this secretive approach to data,” said Francesca Recanatini of the World Bank. “It’s almost like there is a patent on it, and they’re not going to tell you anything about it.” Without understanding how the data was generated and why, there is little to be done to counter any potential inaccuracies in the resulting data or flaws in the methodology. There is also no opportunity for outsiders to learn from innovation or to replicate the approach in different settings.

Although there are recent practices to include local stakeholders in various parts of the project design process, there is still growing fatigue with external assessments in many of the world’s less-powerful countries. Marie Laberge of the UNDP believes that countries are “growing impatient with being assessed from the outside, of being ranked on league tables, and therefore there’s this real desire to be sovereign with regards to data.” In one response, several national statistics offices in Africa have come together under the Strategy for the Harmonization of Statistics in Africa (SHaSA) to establish a standardized methodology for the measurement of governance, peace, and security across the continent. There is almost no interference from the donor community, a fact that has generated considerable momentum within African governments.

Major lessons learned by practitioners in measuring corruption and anti-corruption (summary of main observations from Chapter 4):

1. Always be sure of the purpose of the dataset and what it is designed to do.
2. Data should not be produced in a vacuum. To produce reliable, useful and rigorous data, national stakeholders’ engagement is essential.
3. Data is not the end of the story. In order to have meaningful impact, data should be produced to facilitate both citizen participation and government accountability.
4. There is still a need for global, comparative data to track global trends (e.g., GDP data). The issue is: Could anti-corruption practitioners agree on a global standard methodology?
5. Recognize the complementarity (not comparability) of different kinds of measurements, given that different data are produced for different purposes.
6. To be useful for reform efforts, ensure that data is credible without complex statistical techniques. The buy-in for measurement initiatives is very important.

“This is entirely Africa-led and Africa-owned. The community of African statisticians feels that these are their instruments and that they’re doing the data collection work for their own benefit, not to inform any donor programming or reporting. This is really for African governments to use.”

– Marie Laberge, UNDP Senegal

Data produces standards, whether implicit or explicitly stated. But as noted by Alejandro González Arreola of Gestión Social y Cooperación in Mexico, “each country has its own metrics, its own ideas of how it should be measured and how well it is doing in its transparency and anti-corruption measures.” It is important that governments and NGOs located in less-powerful countries have the space to produce reliable, rigorous data about activities in their own contexts. Moreover, excluding the governments, NGOs and citizens from participating in the *thinking* and *design* of external data-collection projects facilitates a vacuum of power regarding actions and solutions. Kate Dyer, Programme Director of Accountability in Tanzania, believes that “this is not about building society’s capacity to address the real, threatening issues surrounding corruption,” which should be the ultimate goal of outsider involvement in these issues.

The Strategy for the Harmonization of Statistics in Africa (SHaSA)

SHaSA is an entirely African-owned and directed initiative to harmonize statistical data across the continent. The idea of SHaSA is rooted in the belief that to achieve African integration, a common information base is necessary, which can only be produced with harmonized statistics. The initiative is anchored within National Statistics Offices, who are leading the effort to develop two separate instruments to measure 14 priority areas, including governance, security and peace. One instrument is a survey module that can be attached to any ongoing survey done by national statistical offices, and the other is an instrument to connect administrative statistics from various ministries, parliament, the national anti-corruption commission, in essence, all state bodies that could possibly be involved in data production in these topical areas.

Typical survey items include: perception with regards to different institutions involved in corruption, experience on bribery, knowledge of the existence of an anti-corruption body, awareness of whether the government is actually making an effort to address these issues.

Typical administrative data items include: number of corruption cases reported by citizens to a dedicated corruption-reporting mechanism, annual budget of the national anti-corruption bureau per capita, number of corruption cases taken to court relative to the number of investigations completed by the national anti-corruption commission.

The process of building the instruments involved the potential users of this data from the outset. The initiative was started through national validation workshops in each piloting country, with various ministries and agencies that were either producers of data or potential users of this data. Over the course of two days, these various government representatives reviewed the instruments question-by-question, and indicator-by-indicator, and tweaked them in a way that would make better sense in their country. There are core-indicators that need to be preserved continent to continent, but ministries added their own information priorities and needs to complement the core-instruments with their own specifics.

Source: United Nations Economic Commission for Africa, www.uneca.org

There is also the question of real-world relevance and project sustainability in the absence of local involvement in the data-production process. As Brendan Halloran of Transparency and Accountability Initiative expressed, “I’m not sure about the ownership and how these things are actually ingrained in these local communities, or if they’re sort of just flowing in and dropped. Is that really sustainable?” The demand for information is exceptionally important for the success of measurement initiatives. There may be any number of constraints on demand, and external data producers must be careful to recognize the limits of their data for the goal of social change.

“Why do you create an index if nobody is going to use it? Nobody is complaining, nobody has the space for advocacy, nobody can understand the data that you produce. The measurement only satisfies donors or international bureaucracy.”

- Alejandro González Arreola of Gestión Social y Cooperación in Mexico

It may be that the first step is to creating the civic space for these measurements to be useful. Local stakeholders play a significant role in campaigning, lobbying, and generally making use of data in reform efforts. Before they can use or produce data, they need the freedom and interest to advocate for change. Jonathan Fox of American University notes that both internal and external assessments are “simply generating the information as a tool or ammunition for the national civil society actors to guide their efforts.” Alejandro González Arreola suggests that producers ask themselves, “Who will actually use it, for what purpose, and what are they going to get out of it at the end of the day?”

Data is Not the End of the Story

Measurement initiatives on corruption, transparency, accountability, and integrity are not conducted simply to produce data. Practitioners repeatedly stress that their objectives with measurement of any kind is about change, whether eliminating illegal practices, facilitating organizational learning, or improving the lives of citizens.

“There is a trend in the third sector of thinking that we’re all managing projects. We are not. We are doing things to impact people’s lives. So our job is not to sit at a computer and look at a budget in isolation. We have a responsibility to people. We are civil society organizations. We have to make sure that we make the most effort possible to make things work.”

- Rute Caldeira, Transparency International Secretariat in Berlin

It is clear from practitioner insights that “data for accountability” is of paramount importance to the success of any initiative that involves measurement.

Africa Peer Review Mechanisms (APRM)

The APRM is an example of a mutually agreed program, voluntarily adopted by the member states of the African Union, to promote and re-enforce high standards of governance. Evidence of its impact through the use of peer-reviewed assessment includes:

Diagnostic Value of the APRM

Recent developments in reviewed countries have demonstrated the strength of the mechanism as an early warning system for emerging issues and potential crisis. Other country reports that have been made publicly available also highlight common challenges across the continent, including managing diversity, curbing corruption, and strengthening accountability institutions. The diagnostic strength of the mechanism makes it a promising tool to identify key areas of intervention and set priorities for reform.

Governance Gains

As a direct outcome of the review process, Ghana reduced the size of the Cabinet and passed a long-awaited bill to protect whistleblowers and promote access to information. Rwanda reformed its business environment and various governance indicators indicate progress made in terms of control of corruption, government effectiveness and transparency of the regulatory frameworks. Kenya also passed laws on witness protection and public procurement following the completion of the review process.

Development gains

Beyond governance issues, evaluation of the APRM suggests positive outcomes towards supporting the achievement of development goals. It makes a useful contribution to ensuring respect for international commitments. Even if not entirely achieved in practice, the level of citizen participation in the review process comes closer to the ideals of right to development criteria in terms of process than any other similar mechanism.

Source: Marie Chêne, The African Peer Review Mechanism (APRM), U4 Helpdesk and Transparency International, 2009.

Credibility of approach is determined by the use of data to facilitate both citizen participation and government accountability.

“We worked a lot in producing data showing the failings of public policy and presenting the data to authorities. Then little or nothing happens. So we understood that the issue was not really about producing evidence about the fairness of public policy, but trying to implement processes that will involve the citizens affected. It was not that authorities didn’t know the failings. It’s that they actually were not giving any priority to the issue without citizen engagement for accountability.”

- Walter Flores, Centro de Estudios para la Equidad y Gobernanza de los Sistemas de Salud, Guatemala

New forms of mobile platforms for data collection and analysis are changing the way measurement is designed and applied. In places with high penetration of the Internet and mobile technologies, and coupled with an evolving understanding of service delivery as a collaborative endeavor, the focus of policymaking has become much more citizen-centered. Governments are coming to the realization that Web-based citizen reporting on policy and operational failures can inform solutions that ordinarily would have taken much longer to devise.

Be Responsible!

Open Ideas Montenegro was an open competition for the best Web mobile app that addressed critical social issues. It was an experiment in social innovation for the development of tech-powered solutions. Citizens in Montenegro are now equipped with the winning idea – a new mobile app “Be Responsible,” to help them transform them into vigilant reporters, scanning the country for illegal waste dumps, misuse of official vehicles, irregular parking, roadblocks, and failure to comply with tax regulations. The app was developed by several teachers, current and former students of the University of Montenegro’s Faculty of Electrical Engineering.

Within a month of the app launch, the ministry of finance realized the importance of thousands of citizen reports on forms of tax evasion, as it was grappling with a gray, informal economy. The government made a commitment that all reports would be published on the Web, as well as the outcome, so inspection proceedings are being communicated every day with updated statistics. There was a commitment from the government that for each involved fine, where there was a violation of tax regulations or some other laws, a fine would be levied, and 50 percent of the fines would be allocated to community projects chosen by citizens. The team also managed to partner with the Public Relations Bureau of the Government of Montenegro on responding to citizen reports about the misuse of official vehicles.

Source: UNDP, <http://europeandcis.undp.org/>

But technology can be employed successfully in rural areas that face service-delivery failures on a regular basis, and where citizen monitoring is one of the few means of ensuring effectiveness in government-provided services.

Cellphone Text Messages on Service Delivery Failures in Rural Areas

In Guatemala, a new SMS-based platform will be launched by Centro de Estudios para la Equidad y Gobernanza en los Sistemas de Salud (CEGSS) to receive the kind of reports that communities used to have to produce with pen and paper: a lack of drugs, health person not showing up to work, or lack of medical equipment. It is a public access platform that will keep track of reports—it has geo referencing information about where exactly the report is coming from. It will also track the number of days it takes the authorities to respond to the complaint. The platform has the capacity to send complaints immediately through email to authorities at the local level, the provincial and national level.

There will be a short, six-month trial to look at whether there is a different response from authorities when the local authority knows that the same report has been sent to higher-level authorities. And also to see whether a different behavior is elicited in authorities when they know that this is a public platform. Journalists will also be trained and have access to the platforms, so they can monitor the performance of authorities regarding the complaints coming from rural areas.

Source: Centro de Estudios para la Equidad y Gobernanza en los Sistemas de Salud (CEGSS), www.cegss.org.gt

Focus on feedback loops is placing emphasis on the role of citizens in evaluating and observing government activities on a local level. Many of these initiatives are utilizing Internet technology to facilitate real-time feedback, but in cases where Internet penetration is low, other methods such as Social Audit are employed.

Type A: Principal-Initiated and Managed Feedback Systems (Elected Officials)

- “Government Asks” is a multichannel (web, SMS and offline) mechanism to crowdsource policy solutions that has been running in the Brazilian state of Rio Grande do Sul. 360,000 citizens have voted on 36,000 policy proposals drafted by citizens themselves.
- The “Citizen Feedback Model” of Punjab province in Pakistan contacts citizens proactively (<http://www.punjabmodel.gov.pk/>) by SMS after they have used a public service to rate their satisfaction and check whether bribes were requested. The information is relayed to a unit that reports directly to the Chief Minister.
- “Hello Sarkar” (“Hello Government”) in Nepal is a reactive complaints mechanism (<http://nepalofficers.blogspot.co.uk/2011/11/hello-sarkar-hello-government.html>) using widely publicized hotline, SMS numbers and email so citizens can raise problems and complaints with government.
- The World Bank’s own Presidential Delivery Unit dashboard shares key performance information with all the World Bank’s stakeholders: <http://pdu.worldbank.org/sites/pdu2/en/about/PDU>.

Type B: Manager-initiated and Owned Feedback Systems (Project Managers)

- The World Bank’s OnTrack platform supports SMS and online feedback loops between citizens, government, NGOs, implementing agencies and World Bank staff around World Bank-funded projects. It is currently being implemented in Bolivia in the Rural Alliances Project (PAR) and the Bolivia infrastructure programme Barrios de Verdad (PBCV).
- The Department for Rural Development in the State of Andhra Pradesh in India runs one of the biggest and most comprehensive social audit and feedback systems, with 21,000 villages that have been audited 6 times since 2006.
- Uganda’s U-Report SMS-based polling mechanism was developed by Unicef to get real monitoring of social indicators. More than 200,000 U-Reporters have registered to take part across Uganda.

Type C: User-Initiated and Owned Feedback Systems

- CheckMySchool in the Philippines (<http://www.checkmyschool.org/>) has delivered a consistent 50-85% fix rate for hundreds of identified problems across the school system.
- SeeClickFix (<http://seeclickfix.com/>) has contributed to resolving over 1 million problems identified by citizens across over 200 cities in USA and Canada.
- DevelopmentCheck (<http://www.developmentcheck.org/>) is a tracking and reporting tool for Integrity Action’s Community Integrity Building approach, which has improved services for over 4.5 million people in 11 countries.

Source: Frederick Galtung, “The World Bank’s 100% citizen feedback agenda: a daunting challenge and an amazing opportunity,” *Integrity Action*

Practitioners emphasized that citizen-centered accountability mechanisms can be taxing on local communities, particularly poor and marginalized groups, and may not be perceived as worth the effort required. Many time-tested tools that are considered “good practice” are still too abstract in communities where patronage and informal networks are the norm for getting things done. Scorecards may generate discussion with public officials and donors, but may not be accessible for “the communities for whom and about whom this collected scorecard is done, and for whom you really want this information to go to so that they use it in their own demands on public authorities,” as noted by Emmanuel Gyimah-Boadi of the Center for Democratic Development in Ghana. Demand-driven approaches may need to spend time working with communities to create the civic space that underlies citizen-centered accountability.

There is Still a Need for Global, Comparative Data

Corruption measurement first attracted headlines for its global scope and country rankings. These global comparative datasets have an advantage over local, contextual data: they provide an efficient means of summarizing vast amounts of data in order to present a comprehensive macro-level view of a phenomenon. But aggregation can facilitate unfair and inappropriate comparisons across countries, and the data is difficult to interpret for policy-relevant reform. Despite these drawbacks, the field of measurement continues to turn out more indices, on a variety of topics with a diverse range of data, extending from corruption and government performance, to sectorial challenges and general issues of transparency, accountability, and integrity. What has also followed since the backlash against these indices is a more-nuanced understanding of the role of aggregated comparative data. More than ever before, practitioners are aware of the dangers of drawing unwarranted conclusions with the wrong kind of data. And many global data producers make serious efforts to explain the nature of their data, its sources, and its limitations:

“What [global data producers] claim is the idea of having a very basic metric about the very basic elements that are determinants of the variables that they try to assess. None of them argues that they are the perfect measures that can fit every local context and can detail the story of what is happening in every country in the world.”

- Alejandro González Arreola, *Gestión Social y Cooperación, Mexico*

Practitioners also recognize some fundamentally useful applications of this kind of data. Conceptually rigorous aggregated data establishes a comparative framework for discussion and research. Vanessa Tucker of Freedom House proposes that “there is a very important place for the global producers, because you need to have a single methodology for the entire world to look at the world picture.” The most oft-cited need for global comparative data is for global advocacy purposes. Joachim Nahem of the International Law and Policy Institute of Norway notes that this is because “advocacy-wise, they’ve been tremendously successful.” Global, comparative datasets get people to pay attention to corruption, and as Michael Johnston of Colgate University suggests, they are “undeniably useful in terms of agenda-setting and agenda-building.” In fact, Finn Heinrich of Transparency International Secretariat has seen that practitioners value global comparative indices precisely because they “put the topic on the agenda, keep it on the agenda, raise public awareness, get governments interested and open the door to talk.” Global comparative data generates much-needed attention that starts conversations where previously unsolvable and pervasive corruption problems still exist.

Recognize the Complementarity (Not Comparability) of Different Kinds of Measurements

Given the continued need for global comparative data in advocacy efforts, it is important to recognize that not all measurement approaches, nor their results, can be easily compared.

“Numbers and ratings are extremely useful for making sense of very complex processes, to simplify a complex reality. But the numbers have to tell a story. Too often comparisons are reduced to comparing numbers in form of a ranking, for example through direct quantitative comparisons. There are many other, equally relevant and robust comparisons, such as second order comparisons of qualitative information.”

- Finn Heinrich, Transparency International Secretariat, Germany

Measurement approaches are best understood as complementary efforts to capture information about related, and relevant, issues. Global debates spurred by large comparative indices generate discussion at the country level, which may lead to nationally produced datasets on particular issues that are adapting methodologies for local contexts. Practitioners agreed that complementarity is important to building a coherent story around an issue of concern. Finn Heinrich continues by noting that “every measure is different. You have different types of measurements that are all useful, depending on the purpose. The global arena is quite different to the national arena, and to the very local arena.”

The World Justice Project (WJP) Rule of Law Index

This index measures how the rule of law is experienced in everyday life in 99 countries around the globe, based on over 100,000 household and 2,400 expert surveys worldwide. It contains 47 indicators organized around eight themes: constraints on government powers, absence of corruption, open government, fundamental rights, order and security, regulatory enforcement, civil justice, and criminal justice. The **World Justice Challenge** is the WJP’s seed grant competition designed to incubate practical, on-the-ground programs that strengthen the rule of law worldwide. Selected organizations will receive seed grants ranging from US\$10,000-\$20,000 as well as connections to other partners in the WJP network and increased visibility for their work. Applicants use data from the WJP Rule of Law Index to identify rule-of-law weaknesses that could be improved in their communities. Winning projects include:

A2I Toolkit (Global): This program will create an open-source toolkit to support legislative development, monitoring, and compliance exercises on a global scale. The toolkit will be developed using existing materials and an online prototype to rapidly advance several knowledge development goals including comprehensive National, Regional, and International legal frameworks, assessments against established benchmarks, and geo-referencing and map analytics to show patterns and relationships.

Open County Government (Kenya): A new government was elected in 2013 and tasked with implementing the constitution enacted in 2010 which includes focusing more political and economic resources into the county level. This project will help county citizens to engage their leaders on issues of resource allocation, map community assets, and conduct “asset mapping” including village mapping, village street naming, and uploading GIS village maps onto Google maps.

Participlan (Argentina, Bolivia): The project seeks to upgrade and prevent informal settlements through participatory territorial planning, engaging slum dwellers and public officials in a productive dialogue to respond to the challenge of informal settlement expansion.

Source: World Justice Project, <http://www.worldjusticeproject.org/>

Methodologies are complementary in the sense that they capture different kinds of information with a variety of methods, and they can be combined to better capture many sides of the issue being studied. Some kinds of methodologies, such as community scorecards and social audits, generate non-survey qualitative data that cannot be collected across an entire community, city or country. But Christopher Wilson of the engine room notes that “one of the easiest ways for policymakers to dismiss evidence is by pointing to lack of representativeness.” In these cases, complementary approaches can serve to reinforce valid and reliable findings. Even though citizen-generated, ad hoc data may not be representative, it can play an important role in validating the findings of official surveys, and providing more in-depth information, stories, and reports that are missed out in formal data-collection methods. The data can also serve as a triggering mechanism for official audits and investigations. But there is still work to be done to facilitate constructive dialogue between disparate groups of data producers, particularly, as Christopher Wilson has seen, between “the innovative, technology-wielding civil society accountability cohort and government statisticians, about how new kinds of data can complement official statistics and representative surveys.”

I Paid a Bribe (Kenya)

How many reports? 22,480 in two years.

What is done with them? In its first years of existence, I Paid a Bribe did not systematically forward citizen complaints to other institutions. However, reports have been picked up by media outlets and resulted in the arrest and conviction of perpetrators on at least one occasion. Citizens who wish to identify themselves can choose to share their reports with the Central Vigilance Commission (CVC). The CVC is a governmental body mandated to address corruption.

What works? I Paid a Bribe has received extensive media coverage. This helps to build awareness about the tool and attract more citizen reports. Aggregating positive in addition to negative experiences with access to public services can also engage a greater number of people by making participation seem like less of a political act. Highlighting positive reports also rewards public service delivery institutions that refrain from malfeasance.

Challenges: It is difficult to know if the platform reaches new people as opposed to providing a tool for people who are already engaged in civic issues. This is in part because the initiative has until recently required participants to maintain anonymity when reporting. When participants are anonymous, an initiative cannot systematically follow up with them in order to learn about the impact of participation. While the I Paid a Bribe platform provides data analytics and visualizations, this data is likely skewed towards the experiences of more technologically savvy citizens.

Source: New Technologies Against Petty Corruption: Tactics and Lessons from the 2012 International Anti-corruption Conference, the engine room, 2012

There are trade-offs with all measurement approaches, as each approach has its own set of strengths and limitations [see chart in Chapter 2].

“It’s really easy to look at what’s out there right now and say, ‘Oh, that’s all terrible. I’m going to create something that doesn’t have those limitations.’ Well, when it actually comes time to put everything together, it’s nearly impossible. You can’t overcome all of those limitations.”

- Vanessa Tucker, Freedom House, United States

Practitioners recognize the need to be realistic about expectations for measurement initiatives, as there is rarely one solution that will answer every research question. The “generational” framework for understanding measurement categorizes datasets by size and purpose, and serves as one way to think about differences in methodologies. First-generation corruption measurement approaches compiled hundreds of data points using perceptions and experiences as the basis for measurement on a few key issues. Second-generation measurement approaches used multi-dimensional, multi-indicator datasets that compared fewer countries for more focused study with semi-aggregated data. Third-generation measurement approaches are based in only one country or community, often using a consultative approach that digs deeper into core problems. None of these generational approaches is a replacement for another, but instead, they serve different purposes in different contexts.

“Local Data”: Effective at Guiding Reforms, Difficult to Sustain Over Time

Alternatives to global comparative data have proliferated in the last decade, allowing for complementary measurements of corruption and anti-corruption. Practitioners agreed that this lower-level analysis is critical to identifying areas for reform and potential remedying actions, as opposed to global data at the macro level of the international arena.

The Importance of Assessing Local Governance and Corruption

It is important to understand that assessing local governance is not simply a subset or a disaggregated form of national governance assessments. Assessments of local governance provide important information on issues specific to the local level, such as policies vis-à-vis decentralization, participation and local accountability.

Assessments of local governance are undertaken for multiple purposes and reasons:

- To identify potential gaps and constraints in local policy implementation;
- To identify specific capacity development needs and to monitor the results of capacity development efforts;
- To formulate change plans and solicit donor or peer assistance for improving specific aspects of local governance;
- To engage civil society and private sector in local governance; and
- To provide an objective account of achievements of local elected leaders (especially at times of re-elections), and thus building accountability.

One of the main differences between a national and local governance assessment is the greater proximity to the real-world issues. In contrast to national governance, which often deals with systemic policies, the local level is in a daily and intensive interaction with the citizens. Therefore, local assessments need to be much more sensitive to the particular needs of groups of stakeholders and certain segments in the local community.

Source: Wilde et al., A Users' Guide to Measuring Local Governance, UNDP, 2009.

The range of data-collection methods and types of data can be applied in a variety of local contexts; they are suited for any level of analysis, as long as constraints are properly considered. The sphere of "local data," however, encompasses at least two different levels of analysis.

- A common understanding of local consists of micro-level data that concerns *individuals and households*, and is associated with the collection of data through exit surveys, household surveys, or crowdsourcing.
- Another level of local data is concerned with *communities, organizations or agencies, and provinces or states*. This level of data is referred to as meso-level data, situated between macro and micro-levels of analysis. It is associated with hybrid approaches that combine various types of data, such as community scorecards, subnational assessments, and indicator-driven case studies of public-sector functions.

Many of the newer measurement approaches focused on corruption and related issues are oriented to this meso-level analysis, which includes national-level oversight functions, performance of government agencies, and resulting community-level outcomes. As Jonathan Fox of American University suggests, "Think meso, that's my takeaway slogan. Think meso to be useful and tractable." By collecting data on these mid-level practices and outcomes, meso-level data can provide actionable insight that guides reforms in the public sector.

Social Audits of Andhra Pradesh

In 2006 the government of Andhra Pradesh, India mandated the use of social audits for the implementation of the Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS). Social audits are considered low-cost and powerful participatory tools that can bolster awareness and improve public service delivery. A key policy question is the impacts of these audits so far.

A team of researchers investigated the citizen experiences of the social audits in Andhra Pradesh. They found that audits provided a channel of communication through which citizens could interact with government and voice complaints. The audits also raised awareness about entitlements, and shifted the power dynamic between beneficiaries and officials. But another team of researchers found an insignificant impact of audits on employment and a modest decline in the wage leakage. Repeated audits did not deter irregularities. The findings suggest a changing anatomy of corruption, where transgressors keep one step ahead of auditors and respond to more intense scrutiny by locating new avenues for rent extraction.

Clearly, mechanisms like social audits are a powerful tool. But for social audits to result in a responsive and accountable government, they need to be supported by a political commitment to reform the administrative system and challenging political arrangements that support corruption at the local level.

Source: Afridi and Iversen, "Social Audits and Mgnrega Delivery: Lessons from Andhra Pradesh," 2014; Yamini Ayar, "MGNREGS social audit lessons from AP," 2014: <http://www.livemint.com>

Lower level micro and meso-level data provides more accurate information about local contexts, and is responsive to the interests of local stakeholders. It is complementary to global, comparative efforts by following up on findings, digging into what Vanessa Herringshaw of the Transparency and Accountability Initiative calls the "nitty-gritty, on-the-ground issues" that become apparent once the general problems have been identified and acknowledged by governments. As Marie Laberge of UNDP Senegal notes, "Civil society-led assessments that zoom in on a particular sector or a particular institution will be all the more needed once the spotlight has been put on a certain issue at a superficial level." Lower level data also captures information on instances of citizen-government interaction, focusing attention on areas of petty corruption and poor service delivery.

"Most of these global indices tend to measure national issues, but actually it's at the subnational level that it really matters to people. This is where they interact with government and other types of power holders on a daily basis, and that's again much harder to measure, more expensive, etc. But it is extremely valuable."

- Blair Glencorse, Accountability Lab, United States

Corruption and anti-corruption phenomena at this organizational and community level cannot be reliably captured with global expert assessments. In these cases, it is citizen experiences, administrative data, and community-level assessments that are needed to fully understand the kinds of problems and their possible solutions.

But data at the micro and meso-levels has drawbacks. It is often not replicable because it is tailored to specific contexts, and its labor and resource-intensive nature make it difficult to sustain regular data collection over time. It also doesn't generate widespread media attention. Michael Johnston of Colgate University notes the inherent trade-off with policy-relevant data, which is that "it's got to be local, but if it's local it's not going to generate headlines."

Be Careful of Negative Benchmarking and Gaming

Crowdsourcing is an innovative mobile method for capturing experiences and perceptions of various issues, including corruption and related issues. But as noted by Finn Heinrich of Transparency International Secretariat, “It’s not something that you can pick up off the shelf. It requires a good understanding of people’s motivation to participate, and a good understanding of how corruption actually works.” Simply aggregating data without deeper analysis may conflate complaints about bad management with instances of petty corruption.

There is also the possibility that lower-level specific data will drive behaviors in the wrong direction.

“There’s a Vietnamese attempt at an equivalent of Ipaidabribe.com, but it has degenerated into a bulletin board saying you can get your driver’s license for a smaller bribe over here. What does it tell you about crowdsourcing when the crowd is not oriented in the sort of direction we assume they would be?”

- Michael Johnston, Colgate University, United States

One drawback of publicly available, lower-level, specific data is that individuals can easily adjust their behavior if they understand how to undermine the measures. This is referred to as “gaming the system.” Gaming the system occurs when individuals manipulate rules and procedures in order to influence the outcome for their own advantage. As Mihaly Fazekas of the University of Cambridge noted, “We strive for specificity but the real advantage of being not specific is that there’s no way to avoid measurement. If it’s specific, you can avoid it.” Using complementary sources of data is one option for validating findings, but there are clearly trade-offs in every approach.

Ensure that Data is Credible, without Complex Statistical Techniques

Without credibility, data will quickly fade away, despite any level of innovation or media publicity. In fact, lower-level datasets may not even be used without being able to demonstrate that their measures are unbiased, rigorously designed, and based on valid and reliable data. There are a variety of ways to build credibility into a measurement initiative, without relying on complex statistical techniques that may not be accessible to non-experts.

By involving governments as advisers on projects, data producers can facilitate buy-in from traditionally tough critics of data-collection efforts. Data, no matter if it is external or internally generated, often highlight the lackluster performance of government agencies, and can be consequently rebuffed by the people who would most likely benefit from it.

“Well, [the data are] credible with the population and the media, but whether they’re credible with the government is another story. In one sense the government prefers to believe its own data. And there certainly are some people in the government that do take anti-corruption seriously that are trying to work on it; they just have a really hard time using external data.”

- Andrew Wells-Dang, Oxfam, Vietnam

Likewise, national statistics offices can establish the credibility and legitimacy of the data-collection effort.

National statistics offices are much less threatening as an institution. NGO data is perhaps likely to be challenged on methodological grounds, or seen as adverse to the government, but there is confidence between the NSO and the political leadership.

Triangulation of official administrative data with survey data from citizens and the private sector is also a means of establishing credibility in findings. The experience and perceptions of people outside government, combined with public sector performance data, can be used as complementary strengthening mechanisms. Michael Johnston of Colgate University proposes that “this can really be a much more credible and much more comprehensible kind of result,” and it is a relatively straightforward approach to adopt.

Credibility can also be established through rigorous oversight of data collection and analysis efforts. Partnering between local NGOs and donor or research organizations goes both ways to ensuring that data are considered unbiased and reliable. Donors can engage in capacity building with local NGOs, and NGOs can make external data more credible to local stakeholders. But this kind of partnership can help even if local efforts are more than capable of rigorous data collection. As Vanessa Herringshaw of Transparency and Accountability Initiative cautions, “There’s difference between doing it right, and people believing it. Sometimes that only comes through positioning – not through standard-setting.”

In Turn, Data Can Build Credibility in Reform Efforts

Credible data can be used to strengthen reform efforts in a number of areas. In the first instance, it can be used to confirm the experiences of citizens and challenge official denials of existing problems.

“I think most of what these indicators reveal is really well known for the people on the ground. It’s just putting it into a context, giving it kind of the clout of hard science. Then numbers give people some credibility.”

- Mihaly Fazekas, University of Cambridge, United Kingdom

Credible data can also serve as a means of community empowerment. Even though the objective of community scorecards and social accountability practices is not to produce a comparable dataset, what they produce is

“still data, and the value of that is not the data per se, but the processes that you generate around it - the ability of ordinary people, particularly those affected by failing public policies, to produce and analyze data. This generates an ownership of the process by people, which is empowering and at the same time is crucial to sustaining citizens’ demands for accountability.”

- Walter Flores, Centro de Estudios para la Equidad y Gobernanza de los Sistemas de Salud, Guatemala

Managerial empowerment may also be an outcome of credible data. In monitoring and evaluation for accountability purposes, local NGOs are required to demonstrate appropriate use of funding within the donor framework. They are expected to engage in practices that fulfill the objectives of donors, which, according to Kate Dyer of Accountability in Tanzania, can create “difficulty in convincing local NGOs that monitoring can be primarily for them, secondarily for the donor.” Investment in the credibility of the data that is being produced, even if only for monitoring and evaluation purposes, increases the likelihood that it will be taken seriously by others and facilitate organizational learning.

“More importantly, it’s monitoring and evaluation to shape strategic and tactical approaches that are positioned to make a difference on the ground.”

- Brendan Halloran, Transparency and Accountability Initiative, United Kingdom

These kinds of data are needed at all levels of an organization to do planning, budgets, and programming. Credible data allows for better decisions in organizations, increased likelihood of desired outcomes and community-level impacts, and clearer lines of accountability between funders and local organizations.

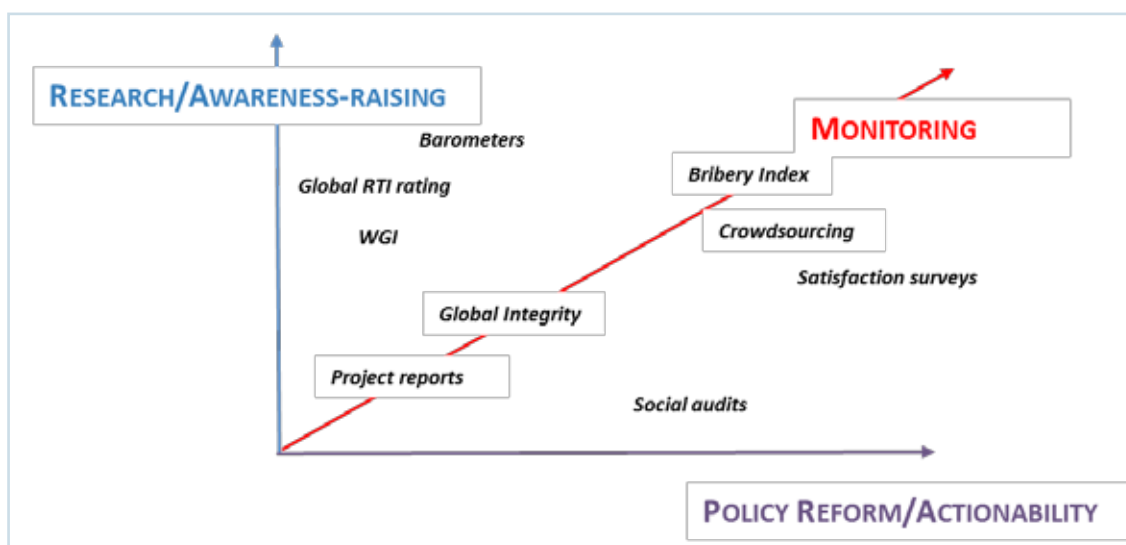
Always Ask – What is the Purpose of the Dataset? What is it Designed to Do?

Purpose is the starting point of any measurement initiative. No matter the tool, or method, or context, data producers must first determine the objective of the data-collection effort. And the purpose, the objective of the data collection, will determine the type of data to gather and methodology to use.

“Think of a three-dimensional space—it’s probably even more but three-dimensional at least—and on one axis you can have research or awareness-raising, and then on another axis you can have policy reform, and on another axis you can have monitoring, and in that space you can place all the different datasets or methodologies. And then you can have a meaningful discussion about what works and why.”

- Francesca Recanatini, World Bank, United States

Figure 18: 3-dimensional Mapping of the Purpose for Measurement Initiatives



Source: : Author and Francesca Recanatini, World Bank

Data producers must also keep in mind that “there is not one tool that will serve all purposes,” noted Finn Heinrich of Transparency International Secretariat. A multi-purpose question might require more than one tool, as Finn Heinrich continues, “particularly in the case of corruption, where most of the existing methods and approaches have limitations.” Moreover, as the level of analysis moves from macro to meso or micro, context becomes even more important for purpose. As Alejandro González Arreola of Gestión Social y Cooperación noted, there is a “serious problem with making blueprints for these types of projects,” because there is no one-size-fits-all approach.

Citizen-generated Data: No Blueprint Approach

Data generated by citizens at the meso-level of communities or micro-level of individuals serve a variety of purposes depending on local needs and demands. More than one approach, tool, design, and method are used sometimes simultaneously to supplement government-administered data.

- Citizen reports on corruption, which may be reported through multiple media, and which may or may not include geographic data, time data, identification of specific institutions or individuals, type of event or financial amounts.
- Citizen perspectives, which may be collected through structured and representative surveys, through exit surveys and points of public service delivery, through online comments to specific themes, legislation or institutions.
- Citizen reports and categorization of suspicious public information, such as procurement announcements or contracts.
- Citizen monitoring of project and public work completion against public budget allocations, to identify instances of institutional corruption.
- Crowdsourced information on public figures, including assets, activities and relationships.
- Unstructured data, such as audio-visual recordings, images and narrative testimonies, which can often be powerful in local campaigning.

Source: CIVICUS support to people-powered accountability and the data revolution: a scoping study by the engine room, April 2014, <https://www.theengineroom.org>

The research community has an important role to play in helping practitioners define success and impact, and developing approaches to measuring those kinds of results. But there is a need to be critical about researcher engagement, because they have their own sets of incentives and potential biases that may hinder organizations ability to learn.

“To be honest, there’s a real disconnect between—and I can see it very clearly—how policy people would look at what the majority of people who measure corruption do as saying ‘okay, that’s interesting, but how can we use this?’ such as for assessments and benchmarks.”

- Nicholas Charron, University of Gothenburg, Sweden

Data users must be aware of the limitations of data, as not all data is appropriate to all applications. For example, global indices have margins of error that may limit simple comparison, and perceptions aren't appropriate for tracking of reform. Alejandro González Arreola of Gestión Social y Cooperación suggests that for both users and producers, "There is no magic bullet."

Be Precise, Even with Fuzzy Concepts

Corruption and anti-corruption are broad concepts that defy standard definition. Everyone sees a different side of the issue, and individuals have different understandings of how the parts fit together.

"We're talking about things like corruption, transparency, accountability, integrity that aren't neat concepts or finite things that we see in other areas of research. How many vaccines are delivered? How many cases of a particular disease occur? These are hard numbers. It's much more difficult to get hard data in this area, so there is a level of judgment."

- Vanessa Tucker, Freedom House, United States

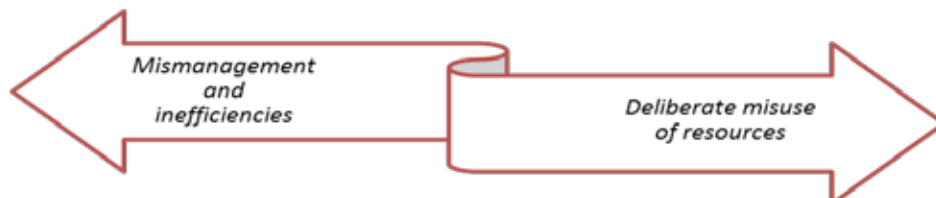
Even in so-called 'objective' measures, there is an element of subjectivity in the choice of topic and selection of cases. In order to avoid confusion in measurement initiatives, it is imperative to use very precise definitions of what you are measuring.

"Our focus is firstly, definitional, so when we talk about corruption it doesn't really make sense to talk about corruption in general, but it makes sense to talk about different kinds of corruption in different kinds of settings."

-Mihaly Fazekas, University of Cambridge, United Kingdom

Often practitioners get caught up in definitions for concepts that simply need to be broken down into their component parts. This can sometimes be accomplished through data disaggregation for gender, age, region, ethnicity, rural vs. urban, etc. But often it must be considered much earlier in the planning process. Corruption and governance are often seen on a continuum, whereby weak governance is considered to be mismanagement and inefficiencies, and corruption is seen as the deliberate misuse of resources. There may be some overlap in efforts for prevention, but sanctions and enforcement for malpractice are most associated with corruption. Despite conceptual challenges, measurement efforts have to clearly define the scope of inquiry, so that it is very clear what is being evaluated.

Figure 11: Governance – Corruption Continuum



Source: Author

Corruption is also often seen as a symptom of lack of accountability and transparency, confusing the cause/effect relationship.

"I think in some cases they confuse symptoms with causes in lots of ways. Again, corruption, for example, 'Okay, let's try and measure it,' but I'm not sure that's actually the cause of the problem. It's just a symptom."

- Blair Glencorse, Accountability Lab, United States

The prevalence of accountability and transparency are viewed as incompatible with corruption, as their existence prevents opportunities for corruption to occur. As Blair Glencorse says, "It's about accountability broadly, and that's how we're going to create system change. It's not about targeting corruption and coming up with technical solutions to what are actually deeply and inherently political problems." This belief has implications for how to interpret findings and propose solutions, but data producers must still be very precise in their definitions of what is being studied.

Beware of Easy Answers

As succinctly stated by Finn Heinrich of Transparency International Secretariat, "Corruption is not an easy phenomenon to detect, or to then solve." All agreed that some of the most daunting questions facing practitioners are how to establish impact, measure outcomes, and track progress. The most widely known measures of corruption over time are global indices of perceptions, and yet, attempting to use this data to track change can quickly derail anti-corruption efforts.

"It raises the possibility that no good ever goes unpunished when you get serious about corruption, because you make a lot of headlines, and the perception variable in Worldwide Governance Indicator rises."

- Michael Johnston, Colgate University, United States

Perceptions can be fickle, and data will quickly reflect the 'sunshine effect,' whereby efforts to curb corruption end up revealing previously hidden illegal practices, and consequently the perceived level of corruption increases. But even with different types of data, and different methodologies, it is exceptionally difficult to capture change when dealing with the diffuse nature and difficult-to-define phenomena associated with corruption, transparency, accountability, and integrity.

There are shifting dynamics of change in the politically charged contexts that characterize corruption and anti-corruption. Technical approaches to monitoring and evaluation interventions often fail because of the nature of accountability and transparency. Blair Glencorse of Accountability Lab suggests that "the context is so important, and the dynamics can change so quickly, that either it's not going to quite measure the right things or it's going to be out of date before it's produced, and so on." Actors must be sensitive to context, making actions very much context-dependent and somewhat unpredictable from a traditional evaluation standpoint.

"Sometimes a little push makes a big ripple when you're strategic and taking advantage of an opportunity, whereas the big push can have little to no effect if it's not strategically oriented where there's a possibility for change."

- Brendan Halloran, Transparency and Accountability Initiative, United Kingdom

There are multiple actors, multiple challenges, and multiple links in the results chain. Changing behavior means changing attitudes and incentive structures, sometimes at different points in time.

“So it’s a collective action story, and you need to act on many grounds at the same time, many fields at the same time. It doesn’t work to say, ‘Ah, first establish an independent judiciary, and then reform political party finances.’ People in powerful positions know the rules of the game, and they will prevent you from building an independent judiciary. It’s really about power rather than just rules.”

-Mihaly Fazekas, University of Cambridge, United Kingdom

Rather than focusing on the trajectory of project activities, it may be more beneficial to consider the factors that contribute to change and how they interact with each other. Timing becomes an important constraint for effective tracking of progress. Change may happen years in the future and/or depend on unexpected windows of opportunity. A very linear time-bound approach may not be appropriate. Many practitioners agree with Finn Heinrich of Transparency International Secretariat, in believing that “non-linear approaches to tracking impact might be more useful. The more traditional outcome monitoring frameworks don’t work – what could work?”

There is a prevailing presumption that failures can be attributed to the tools, especially when innovation is often so tool-focused.

“If you’re going to evaluate how good a measurement tool is, consider if there is something built in, such as whether resources are assigned, or if there is a logical mechanism that can say something about impact tracking. There aren’t usually a lot of incentives to do that. It’s more about the tool and innovation and gathering the data.”

- Joachim Nahem, International Law and Policy Institute, Norway

But failure can appear at any point in the measurement process or project implementation: design, data collection, analysis, etc. Projects must engage in what Brendan Halloran of Transparency and Accountability Initiative calls a “culture of critical inquiry.” He suggests that there should be “opportunities for strategic reflection, and...flexibility in approaches...that will allow you to adapt and integrate based on what you’re finding out about what’s working and what’s not.”

There is no question that capturing impact is a critical piece of the accountability puzzle. But like Vanessa Herringshaw of Transparency and Accountability Initiative, practitioners are genuinely worried that randomized controlled trials with standardized interventions “do not easily lend themselves to governance-type interventions, which need to be flexible and illustrative, complex, evolving.”

Ultimately, there is a tension between adaptive learning and generalizable findings. Many practitioners like Brendan Halloran would assign “primacy to building the learning and adaptive capacity of organizations to work within their context,” rather than privileging the aim of capturing data on outcomes and impact.

Good Practices for Accountability

This chapter presents a three-part definition of ‘good practices for accountability’ in measurement initiatives. Several award-winning examples of innovative measurement initiatives are presented, along with a discussion of complementarity, knowledge-building, and sustainability.

Good practices for accountability **use resources appropriately, efficiently, and inclusively** with the goal of bringing about changes in corruption. In measurement initiatives, good practices vary according to context and purpose. But they require informed action – practitioners must familiarize themselves with the types of data, methods, and methodologies available to best suit their needs. They must also understand the limitations of each approach so that data is not biased or used inappropriately to assess countries or institutions.

Good practices also **incorporate stakeholders in the measurement initiative**, at any or all points in the process, from design to dissemination. Voices of those most affected by the project must be included to ensure that project design is appropriate and targeted. The risks to stakeholders must be carefully weighed so that no retaliation happens in response to collection or dissemination of controversial or highly politicized results, particularly in sensitive cases of corruption.

Finally, good practices **link data to learning and accountability**, as data should always be a means to an end. Data should never “sit on a shelf and collect dust.” It is meant to foster awareness, but more importantly, its primary aim should be to facilitate behavioral change that improves the quality of life for citizens. Feedback mechanisms that collect citizen observations and experiences are an important example of this type of link between data and accountability.

Spotlight on Good Practices

The following measurement initiatives employ good practices for accountability as described above. Several of them have either won or been nominated for awards for innovative practices. They use methodologies appropriately or develop new forms of measurements based on the results of previous interventions. They also incorporate stakeholder participation as an integral feature in design, data collection, and analysis. Finally, they intend for data to be used as a means of enhancing transparency, accountability, and integrity of government, or as a means of improving future measurement initiatives.

(1) Participatory Action Research and Monitoring of Health Services by Marginalized Communities, Guatemala

Using a rights-based approach to community monitoring of public health care services, a civil society organization (CEGSS), in partnership with Community Based Organizations (CBOs), has developed and implemented a system to implement citizens' vigilance of public health care services. The approach includes three main components: a) capacity building process aimed to enhance knowledge and skills of indigenous populations about the legal framework, human rights and public policies and services b) data collection and analysis tools to monitor public services used by the communities themselves and c) strategic advocacy to influence public policy making and services. After 5 years of implementation, communities have influenced the improvement of public health services at local level and the allocation of resources. A formative evaluation conducted in 2010 identified several health and governance outcomes: more personnel in certain health services, sanctions for doctors for poor treatment of indigenous people, better hours for provision of services, services during the weekend, and awareness among health staff that problems exist and improvements are possible. However, the most important achievement has been the connection of highly marginalized communities with public policies and services and the overall Estate in Guatemala. This is resulting in the political empowerment of community leaders that were alienated from public institutions before. At the same time, public health care services in rural areas are receiving an increased attention and resources from central government due to the communities' monitoring and demands for improvements. All of this is resulting in an increment in the level of trust and collaboration between public authorities and CBOs.

Source: Walter Flores, Political empowerment of marginalized indigenous communities through the monitoring of public health care services in Guatemala, Centro de Estudios para la Equidad y Gobernanza de los Sistemas de Salud (CEGSS), 2012.

(2) Citizen Participatory Audits, Philippines

The Commission on Audit's Citizen Participatory Audit (CPA) Project aims to increase awareness and government transparency through citizen involvement in the audit process. It is a constructive and collaborative partnership between the government and citizens through civil society organizations (CSOs). The partnership has developed a mix of tools such as surveys, scorecards, and data-gathering activities that involve citizens in the technical process of COA auditing. It is composed of CSOs, students, private sector and academic groups that conduct joint audits of selected government projects in conjunction with the Commission on Audit. This has resulted in the implementation of three participatory audits to assess the effectiveness of conditional cash transfers (CCTs), a solid waste management program, and a flood control project. The output of the social audit bears the weight of COA's authority, as it becomes part of the official reports. Local CSO partners may also disseminate the findings of the audit reports through radio programs and community meetings. The CPA illustrates how to enhance external scrutiny through citizen engagement in the audit process, though impact depends on the capacity of the COA to compel reform in government projects.

Source: Open Government Partnership, www.opengovpartnership.org

(3) Report Corruption, Macedonia

Transparency International Advocacy and Legal Advice Centres (ALACs) in Macedonia created an online reporting platform and used social media to solicit reports of corruption from younger audiences; citizens can also send in reports to the platform using a mobile application that is available for iPhone or Android. Reports are examined by TI Macedonia staff and then converted into cases, which the ALAC staff use in strategic advocacy. Requiring verification of identity is necessary for Report Corruption to work, as ALAC program staff would not be able to advocate on behalf of anonymous reports. ALAC staff add comments to reports on the website to keep the person who sent in the report updated about its status. The support of TI's ALAC allows Report Corruption to conduct systematic follow-up with citizen reporters, creating feedback loops that keep them engaged and build trust. Report statuses are kept up to date on the comments section of the online platform, which is visible to the public. Advocacy impact based on ALAC reports is monitored through TI's impact-assessment process and shared with constituencies and stakeholder communities.

Source: New Technologies Against Petty Corruption: Tactics and Lessons from the 2012 International Anti-corruption Conference, the engine room, 2012. https://www.theengineroom.org/wp-content/uploads/CIVICUS.Scoping.Study_WEB_.pdf

(4) DevelopmentCheck, Various Countries

DevelopmentCheck is a user-driven and solution-oriented online tool for citizen feedback on the transparency, participation and effectiveness of development projects. Community volunteers are trained by international and country partners in the Community Integrity Building approach so they can effectively monitor development projects in their communities. The community identifies priority development projects to monitor and then collects data on these through project site visits, beneficiary surveys, information requests and photos. They enter the data for each project monitored into a questionnaire on DevelopmentCheck. They can also upload project documents such as the bill of quantity or contract if available, as well as photos or videos. Once data is uploaded online, a moderator verifies and publishes the information through DevelopmentCheck. As some projects are monitored over a long period of time data may be entered from more than one site visit. The country partners share their findings with the community, government and contractors. Through constructive engagement with these key stakeholders, they can fix problems and improve development projects. The percentage of problems resolved is called the 'fix-rate,' which is used to track progress. DevelopmentCheck also enables cross-country comparison and collective advocacy. In other words, this data provides evidence to generate pressure for improved services and projects, such as in Timor Leste, where District Monitoring Committees comprise monitors, contractors, community members and local authorities, work together to find solutions or "fixes" to identified problems.

Source: DevelopmentCheck – online citizen feedback on development projects and <http://www.developmentcheck.org/>

(5) Improving School Bursaries, Kenya

In 2008, the Institute for Policy Analysis and Research (IPAR) in Kenya sought to track resource flows for the country's Secondary Education Bursary Fund. Starting in Nairobi province, IPAR identified significant inefficiencies with the scheme, including that 20 percent of schools were receiving bursaries for students no longer enrolled and other evidence of leakage. IPAR proceeded to work with the government and private bursary providers to standardize records and processes for bursary disbursement. They further shared these findings widely to arm schools and citizens with information to understand how the bursary was working in their communities. Early evidence from IPAR suggests that the bursary fund has since reduced many of the inefficiencies that were found in the original 2008 study.

Source: Where Are Our Budgets? Public Expenditure Tracking Surveys as Tools for Engaging with Civil Society, USAID, 2014. <https://www.hfgproject.org/wp-content/uploads/2014/04/PETSCivilSocietyBrief.pdf>

(6) Participatory Monitoring of the State of Forest and REDD+ Governance in Indonesia

Since 2012 the Participatory Governance Assessment for REDD+ in Indonesia, designed and guided by a multi-stakeholder Expert Panel, has allowed the generation of quality and consensual data on the governance of forests and the national REDD+ (reducing emissions from deforestation and forest degradation) process. Anti-corruption is measured through a subset of ten indicators that relate to regulations, civil society actions, capacities and private sector perceptions, and law enforcement capacity through an additional nine indicators. The 2012 edition of the PGA resulted in the "Indonesia Forest and REDD+ Governance Index", and led to among others, in-depth integrity check of the forest permits system (from both user and providers perspectives), the development of sanctions for public officials to comply with integrity standards and encourage non-corrupt practices (led by the anti-corruption commission UKP4), and the establishing of a conflict resolution mechanism. The next edition is expected in late 2015.

Source : PGA Fast Facts, UN-REDD Programme, 2013

Complementarity: Context Matters. Purpose Matters.

Complementarity of approaches matters more to measurement effectiveness than innovation in tools or techniques. Every methodology, tool, or technique has its strengths and limitations, which can vary according to context and purpose. Because context and purpose are fundamentally important to the success of a measurement initiative, replicability of approaches requires fine-tuning to fit the circumstances at hand. No one approach will provide all the answers, nor will one approach work in all contexts.

Similarly, local data at the micro or meso-level answers very different questions to global, comparative data, but lower levels of analysis are often initiated or driven by findings at the macro-level. There is a clear complementary relationship between the levels of analysis that can help practitioners to better understanding the story of corruption or anti-corruption in a single context.

In order to impact policy, or to feed data and information into policy reforms, what matters is not just whether corruption can be measured precisely. Indeed, it is questionable whether this is possible given the multi-dimensionality of corruption and related phenomena. The question then revolves around whether circumstances can be measured effectively with proxy indicators that can track progress and change. Proxy indicators are alternatives to objective indicators that more directly measure the phenomenon under study but that may be hard to operationalize or require costly data collection.⁸⁶ Proxy measurements are commonly used in other sectors to supplement direct measurements. Examples include vaccination rates as proxy measurement for incidence of disease, or household consumption rates for incidence of poverty.

⁸⁶Jesper Johnson and Phil Mason, *The Proxy Challenge: Why bespoke proxy indicators can help solve the anti-corruption measurement problem*, Chr. Michelsen Institute, 2013.

Other areas for potential proxy measurements include:

- Correlation between social trust in public institutions (particularly service-delivery institutions) and changing levels of corruption in those institutions.
- Linkage between levels of tax compliance and corruption levels.
- Number of ghost workers as a proxy indicator for payroll fraud.
- Number of senior civil servants changing jobs after a change in government as an indicator of patronage.

The best strategy may be to avoid spending effort in identifying a gold standard, but rather, making the best out of the existing measurement approaches, and achieving consensus with major stakeholders about their use and misuse.

Knowledge of the Commons

In order for measurement approaches to continue to evolve in more robust and refined ways, it is imperative that practitioners share their knowledge, methods, and data with the public. Perhaps more importantly, practitioners should share their understandings of change and experiences of failure with the community of practice. Collaboration among various stakeholders encourages learning and experimentation, while minimizing the risks involved with innovation.

State-centric monitoring processes make it difficult for civil society to participate or utilize data, particularly when there is no mandatory mechanism for releasing data within short timelines. The UNCAC review mechanism is a clear example of this challenge, as state parties are allowed to maintain confidentiality of the review process and the resulting data. But this is hardly the only example in which data is not released to a broader public, even though it pertains to the public good.

Improving the effectiveness of anti-corruption interventions entails the release of data for accountability and learning purposes. It also involves the documentation of evidence-based successes and lessons learned that are then disseminated to a larger audience. The integration of those lessons into other measurement initiatives and anti-corruption efforts impacts the quality of future interventions, as well as the implementation of systems for tracking change.

Partnership and Sustainability

Lower-level analysis that collects data on individuals, households, communities, organizations, and provinces/states is much more contextually-dependent than global, comparative data. Meeting project objectives of sustainability, credibility, and relevance depends on local stakeholder interest and capacities, as well as government buy-in from local officials.

There are often gaps in coverage and purpose between community-level measurement initiatives and national or global initiatives. Partnership between these levels can be fostered to ensure that data is relevant to a wider range of users.

It is also important to facilitate change through support to local partners, rather than direct intervention. Structures of accountability between donors and “subcontracting” NGOs limit the space for stakeholders to adapt, learn, and thrive on their own terms. Ownership of measurement approaches and resulting data is an important factor in the use of data for accountability purposes. If stakeholders feel that data is irrelevant to their needs, or too abstract for use in policy reform, then data will not be used.

References

- Afridi, F., & Iversen, V. (2014). *Social Audits and Mgnrega Delivery: Lessons from Andhra Pradesh* (SSRN Scholarly Paper No. ID 2424194). Rochester, NY: Social Science Research Network.
- Aiyar, Y. (2014, March 20). MGNREGS social audit lessons from AP. *Livemint*. India. Retrieved from <http://www.livemint.com/Opinion/V3tTIWCHiDjNT2GP45geeN/MGNREGS-social-audit-lessons-from-AP.html>
- Alsop, R., & Heinsohn, N. (2005). *Measuring Empowerment in Practice: Structuring Analysis and Framing Indicators*. Washington, DC: World Bank.
- Arndt, C., & Oman, C. (2006). *Uses and abuses of governance indicators*. OECD Publishing. Retrieved from <http://www.oecd.org/dev/poverty/usesandabusesofgovernanceindicators.htm>
- Besançon, M. (2003). *Good Governance Rankings: The Art of Measurement*. World Peace Foundation.
- Beth, E. (2013). *Strengthening the corruption prevention framework*. OECD Tunisia.
- Blatter, J., & Haverland, M. (2012). *Designing Case Studies: Explanatory Approaches in Small-N Research*. New York, NY: Palgrave Macmillan.
- Blundell, A. G., & Harwell, E. E. (2010). *Analyzing Corruption in the Forestry Sector*. Berlin: Transparency International.
- Campos, J. E., & Pradhan, S. (2007). *The Many Faces of Corruption: Tracking Vulnerabilities at the Sector Level*. Washington, D.C: World Bank Publications.
- Carter Center. (2014). *Women and the Right of Access to Information in Liberia: A Mixed-Methods Study*. Atlanta, GA: The Carter Center. Retrieved from <http://www.cartercenter.org/resources/pdfs/peace/ati/women-and-ati-prepub-07172014.pdf>
- Coffman, J. (2010). *Advocacy Toolkit Companion*. New York: UNICEF.
- Davies, R., & Dart, J. (2005). *The "Most Significant Change" (MSC) Technique: A Guide to Its Use*.
- DFID. (2010). *How To Note: Addressing Corruption in the Health Sector*. London: DFID.
- Donchev, D., & Ujhelyi, G. (2009). *What Do Corruption Indices Measure? SSRN eLibrary*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1124066
- Dyer, K. (2012). *Making "Evidence" the Plural of "Anecdote": A "Work in Progress" using Outcome Mapping and the Logframe in governance and accountability programming in Tanzania*. DFID.
- European Commission. (2013). *Evalsed Sourcebook: Methods and Techniques*.

- Flores, W. (2012). *Political empowerment of marginalized indigenous communities through the monitoring of public health care services in Guatemala*. Centro de Estudios para la Equidad y Gobernanza de los Sistemas de Salud (CEGSS).
- Fox, J. (2014, March 3). *Social Accountability: What does the evidence really say?*. Presented at the Global Partnership for Social Accountability (GPSA), World Bank.
- Fukuda-Parr, S., Guyer, P., & Lawson-Remer, T. (2011). *Does Budget Transparency Lead to Stronger Human Development Outcomes and Commitments to Economic and Social Rights?* (International Budget Partnerships Working Paper No. 4 No. ID 2211584). Rochester, NY: Social Science Research Network.
- Galtung, F. (2006). *Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices*. In C. Sampford, A. Shacklock, C. Connors, & F. Galtung (Eds.), *Measuring corruption* (pp. 101–130). Ashgate Publishing, Ltd.
- Geber, T. (2014, May 29). *Why should we care about comparability in corruption data?* Retrieved from <http://schoolofdata.org/2014/05/29/why-should-we-care-about-comparability-in-corruption-data/>
- Goel, R. K., & Mazhar, U. (2014). *A Replication of "Corruption and Elections An Empirical Study for a Cross-section of Countries"* (Economics and Politics, 2009). *Public Finance Review*, 1091142114537890. doi:10.1177/1091142114537890
- Group of States against Corruption (GRECO). (2012). *Lessons learnt from the three Evaluation Rounds (2000-2010)*.
- Halloran, B. (2014). *Thinking and Working Politically in the Transparency and Accountability Field*. London, England: Transparency and Accountability Initiative.
- Hanna, R., Bishop, S., Nadel, S., Scheffler, G., & Durlacher, K. (2011). *The effectiveness of anti-corruption policy: What has worked, what hasn't, and what we don't know* (Technical Report). London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Hechler, H., Morris, D., Zinkernagel, G. F., & Koechlin, L. (2011). *Can UNCAC address grand corruption? A political economy analysis of the UN Convention against Corruption and its implementation in three countries*. Chr. Michelsen Institute (CMI).
- Heywood, P. M., & Rose, J. (2014). *"Close but no Cigar": the measurement of corruption*. *Journal of Public Policy, FirstView*(Supplement -1), 1–23. doi:10.1017/S0143814X14000099
- Holland, J., & Ruedin, L. (2012). *Monitoring and evaluating empowerment processes*. Swiss Agency for Development and Cooperation.
- Hulme, D., Savoia, A., & Sen, K. (2014). *Governance as a Global Development Goal? Setting, Measuring and Monitoring the Post-2015 Development Agenda* (SSRN Scholarly Paper No. ID 2425238). Rochester, NY: Social Science Research Network.

- Jain, A. K. (2001). *Corruption: A Review*. *Journal of Economic Surveys*, 15(1), 71–121. doi:10.1111/1467-6419.00133
- Johnson, J. (2014). *Cost-effectiveness and cost-benefit analysis of governance and anti-corruption activities*. Bergen: Christian Michelsen Institute.
- Johnson, J. (2012). *Theories of change in anti-corruption work: A tool for programme design and evaluation*. Bergen: Christian Michelsen Institute.
- Johnson, J., Hechler, H., De Sousa, L., & Mathisen, H. (2011). *How to monitor and evaluate anti-corruption agencies: Guidelines for agencies, donors, and evaluators*. Bergen: Christian Michelsen Institute.
- Johnson, J., & Mason, P. (2013). *The Proxy Challenge: Why bespoke proxy indicators can help solve the anti-corruption measurement problem*. Bergen: Chr. Michelsen Institute.
- Johnson, J., Taxell, N., & 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* (Report No. 7/2012). Bergen: Christian Michelsen Institute.
- June, R., Chowdhury, A., Heller, N., & Werve, J. (2008). *A Users' Guide to Measuring Corruption*. Oslo: UNDP Oslo Governance Center. Retrieved from http://www.undp.org/content/dam/aplaws/publication/en/publications/democratic-governance/dg-publications-for-website/a-users-guide-to-measuring-corruption/users_guide_measuring_corruption.pdf
- Kabeer, N. (1994). *Reversed Realities: Gender Hierarchies in Development Thought*. Verso.
- Kaufmann, D., & Kraay, A. (2007). *On Measuring Governance: Framing Issues for Debate*. Issues paper for January 11, 2007, Roundtable on Measuring Governance, World Bank Institute and Development Economics Vice-Presidency, The World Bank, Washington, DC.
- Kaufmann, D., & Kraay, A. (2008). *Governance Indicators: Where Are We, Where Should We Be Going?* *World Bank Research Observer*, 23(1), 1–30.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). *The Worldwide Governance Indicators: Methodology and Analytical Issues*. Washington, DC: The World Bank.
- Klitgaard, R. (1991). *Controlling Corruption* (Reprint edition). Berkeley: University of California Press.
- Knack, S. (2006). *Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators* (Policy Research Working Paper Series No. 3968 No. 3968). Washington, DC: The World Bank. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=923275
- Kraay, A., & Tawara, N. (2010). *Can disaggregated indicators identify governance reform priorities?* (No. 5254). Washington, DC: The World Bank.
- Kumar Mishra, A., & Gupta, N. (2012). *Assessment of Integrity Pact (IP) In IP Compliant Public Sector Undertakings*. Transparency International India.

- Lewis, M., & Pettersson, G. (2009). *Governance in Health Care Delivery: Raising Performance*. Washington DC: World Bank.
- Lin, M.-W., & Yu, C. (2014). *Can Corruption Be Measured? Comparing Global Versus Local Perceptions of Corruption in East and Southeast Asia*. *Journal of Comparative Policy Analysis: Research and Practice*, 16(2), 140–157.
- Martini, M. (2012). *Overview of Integrity Assessment Tools*. Transparency International and CHR Michelsen Institute.
- Matsheza, P., & Timilsina, A. R. (2008). *Anti-Corruption Interventions for Poverty Reduction, Realization of the MDGs and Promoting Sustainable Development*. New York: UNDP.
- McDevitt, A. (2012a). *Gateway: Mapping the Corruption Assessment Landscape*. Transparency International.
- McDevitt, A. (2012b). *Monitoring Conventions Topic Guide*. Berlin: Transparency International.
- Moore, M., & Teskey, G. (2006). *The CAR Framework: Capability, Accountability, Responsiveness. What Do These Terms Mean, Collectively and Individually? A Discussion Note for DFID and Conflict Advisers*. Institute of Development Studies.
- Nadgrodkiewicz, A., Nakagaki, M., & Tomicic, M. (2012). *Improving Public Governance: Closing the Implementation Gap Between Law and Practice*. Washington, DC: Center for International Private Enterprise (CIPE) and Global Integrity.
- Oxfam. (2013). *Process Tracing: Draft Protocol*.
- Pierce, R. (2008). *Research Methods in Politics*. Thousand Oaks, CA: SAGE Publications Ltd.
- Purushothaman, S., Tobin, T., Vissa, S., & Pillai, P. (2012). *Seeing Beyond the State: Grassroots Women's Perspectives on Corruption and Anti-Corruption*. New York: UNDP.
- Schedler, A. (1999). *Conceptualizing Accountability*. In L. J. Diamond, M. F. Plattner, & A. Schedler (Eds.), *The Self-restraining State: Power and Accountability in New Democracies* (pp. 13–28). Boulder, CO: Lynne Rienner Publishers.
- Schouten, C. (2012). *ICT for Open Contracting in Fragile and Conflict-Affected States*. Washington DC: World Bank.
- Scott, C., & Wilde, A. (2006). *Measuring Democratic Governance: A Framework for selecting pro-poor and gender sensitive indicators*. UNDP.
- Scriven, M. (2008). *A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research*. *Journal of MultiDisciplinary Evaluation*, 5(9), 11–24.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the Range of Designs and Methods for Impact Evaluations* (Working Paper 38). London, England: DFID.

- Tamyalew, A. (2010). *A Review of the Effectiveness of the Anti-corruption and Civil Rights Commission of the Republic of Korea*. Washington, DC: World Bank.
- the engine room. (2014). *CIVICUS support to people-powered accountability and the data revolution: a scoping study*. The Engine Room. Retrieved from <https://www.theengineroom.org/wp-content/uploads/CIVICUS.Scoping.Study.WEB.pdf>
- Thomas, M. (2010). *What Do the Worldwide Governance Indicators Measure? European Journal of Development Research*, 22, 31–54.
- Transparency International. (2013). *Integrity Pacts in the Water Sector: An implementation Guide for Government Officials*.
- Trapnell, S. E. (2010). *Commission on the Prevention of Corruption Slovenia: A Review of the Effectiveness of Anti-Corruption Agencies*. Washington, D.C: World Bank.
- Trapnell, S. E. (2011). *Actionable Governance Indicators: Turning Measurement into Reform*. *Hague Journal on the Rule of Law*, 3(02), 317–348.
- UNDP. (2007). *Governance Indicators: A Users' Guide (2nd Edition)* (p. 100). Oslo, Norway: United Nations.
- UNDP. (2008). *Mainstreaming Anti-Corruption in Development*.
- UNDP. (2009a). *A Trainer's Manual on Measuring Corruption at the Country Level*, UNDP Oslo Governance Centre, Oslo.
- UNDP. (2009b). *Handbook on Planning, Monitoring and Evaluating for Development Results*. New York.
- UNDP. (2010). *Measuring Capacity*. New York: UNDP.
- UNDP. (2011). *Practitioners' Guide: Capacity Assessment of Anti-Corruption Agencies*. New York: UNDP.
- Van Wicklin III, W. A., & Gurkan, A. (2013). *How-to notes : participatory and third party monitoring in World Bank projects - what can non-state actors do?*. Washington, DC: World Bank.
- Voigt, S. (2009). *How (Not) to Measure Institutions*. SSRN eLibrary. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1336272
- White, H. (2009). *Theory-Based Impact Evaluation: Principles and Practice* (Working Paper). International Initiative for Impact Evaluation.
- Wilde, A., Narang, S., Laberge, M., & Moretto, L. (2010). *A Users' Guide to Measuring Local Governance*. Norway: UNDP Oslo Governance Centre.
- World Bank. (2009). *Monitoring and Evaluation: Some Tools, Methods, and Approaches*. Washington, DC.
- World Bank. (2010). *Building Public Support for Anti-corruption Efforts: Why Anti-corruption Agencies Need to Communicate and How?* Washington, DC.

Annex 1: List of Interviewees

Rute Caldeira	Transparency International Secretariat (TI-S)	Germany
Nicholas Charron	University of Gothenburg	Sweden
Kate Dyer	Accountability in Tanzania	Tanzania
Mihaly Fazekas	University of Cambridge	United Kingdom
Walter Flores	Centro de Estudios para la Equidad y Gobernanza de los Sistemas de Salud (CEGSS)	Guatemala
Jonathan Fox	American University	United States
Blair Glencorse	Accountability Lab	United States
Alejandro González Arreola	Gestión Social y Cooperación (GESOC)	Mexico
Emmanuel Gyimah-Boadi	Center for Democratic Development (CDD)	Ghana
Brendan Halloran	Transparency and Accountability Initiative (TAI)	United Kingdom
Andria Hayes-Birchler	Millennium Challenge Corporation (MCC)	United States
Finn Heinrich	Transparency International Secretariat (TI-S)	Germany
Vanessa Herringshaw	Transparency and Accountability Initiative (TAI)	United Kingdom
Jesper Johnson	U4 Anti-Corruption Resource Centre (CHR. Michelsen Institute)	Norway
Michael Johnston	Colgate University	United States
Marie Laberge	United Nations Development Program (UNDP)	Senegal
Marija Novkovic	United Nations Development Program (UNDP)	Montenegro
Joachim Nahem	International Law and Policy Institute (ILPI)	Norway
Alicia Phillips Mandaville	Millennium Challenge Corporation (MCC)	United States
Francesca Recanatini	The World Bank	United States
Fernando Straface	Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento (CIPPEC)	Argentina
Vanessa Tucker	Freedom House	United States
Andrew Wells-Dang	Oxfam	Vietnam
Christopher Wilson	the engine room	Norway

Annex 2: Datasets

Dataset	Organization	Links
Afrobarometer	Center for Democratic Development Ghana, Institute for Empirical Research in Political Economy Benin, Institute for Development Studies Nairobi, Institute for Justice and Reconciliation in South Africa	www.afrobarometer.org/
Americasbarometer	Latin American Public Opinion Project	http://www.vanderbilt.edu/lapop/
Arab barometer	Arab Reform Initiative	http://www.arabbarometer.org/
Asianbarometer	Institute of Political Science, Academia Sinica and the Institute for the Advanced Studies of Humanities and Social Sciences, National Taiwan University.	www.asianbarometer.org/
Bribe Payers Survey	Transparency International	www.transparency.org/research/bpi/overview
RTI Implementation Assessment Tool	Carter Center for Human Rights	http://www.cartercenter.org/peace/ati/IAT/index.html
Global RTI Rating	Centre for Law and Democracy/AccessInfo Europe	http://www.rti-rating.org/
Global Corruption Barometer	Transparency International	http://www.transparency.org/research/gcb/overview
Corruption Perceptions Index	Transparency International	www.transparency.org/research/cpi/
East African Bribery Index	Transparency International Kenya	http://tikenya.org/index.php/more-links/publications/corruption-surveys/east-africa-bribery-index
Enterprise Surveys	World Bank	http://www.enterprisesurveys.org/
European Commission for the Efficiency of Justice (CEPEJ)	European Commission for the Efficiency of Justice (CEPEJ)	http://www.coe.int/T/dghl/cooperation/cepej/default_en.asp
Fix-rate	Integrity Action	http://www.integrityaction.org/siobhan/fix-rate---citizens-can-make-difference
GAC Diagnostics	World Bank	http://go.worldbank.org/XRN2JAJ180
Global Integrity Report	Global Integrity	www.globalintegrity.org
Reports on the Observance of Standards and Codes	IMF	http://www.imf.org/external/NP/rosc/rosc.aspx
Open Budget Index	International Budget Partnership	internationalbudget.org/what-we-do/open-budget-survey
Kenya Bribery Index	Transparency International Kenya	www.tikenya.org/index.php/kenya-bribery-index

Annex 2: Datasets *(continued)*

Dataset	Organization	Links
Latinobarometer	Latinobarómetro Corporation	http://www.latinobarometro.org/lat.jsp
Government at a glance : Transparency in Governance	OECD	http://www.oecd.org/gov/governmentataglance2011.htm
Methodology for Assessing Procurement Systems (MAPS)	OECD	http://www.oecd.org/development/effectiveness/commonbenchmarkingandassessmentmethodologyforpublicprocurementsystemsversion4.htm
Open Data Index	Open Knowledge Foundation	https://index.okfn.org/
Public Expenditure and Financial Accountability (PEFA)	World Bank	http://www.pefa.org
Public Expenditure Tracking Surveys	World Bank	http://pets.prognoz.com/prod/
Resource Governance Index	Revenue Watch Institute	http://www.resourcegovernance.org/rgi
Tax Administration Diagnostic Assessment Tool (TADAT)	IMF	http://www.tadat.org/
National Integrity Systems Assessment	Transparency International	http://www.transparency.org/whatwedo/nis
UNCAC Review Mechanism Reports	UNODC	https://www.unodc.org/unodc/en/treaties/CAC/IRG.html
Crime Statistics	UNODC	https://www.unodc.org/unodc/en/data-and-analysis/crimedata.html
Public Accountability Mechanisms	World Bank	http://www.agidata.org/pam
Rule of Law Index	World Justice Project	http://worldjusticeproject.org/rule-of-law-index
Open Data Barometer	World Wide Web Foundation	http://www.opendataresearch.org/barometer

Annex 3: Experimental and Quasi-experimental Methods Source: Poverty Action Lab

	Methodology	Description	Who is in the comparison group?	Required Assumptions	Required Data
Quasi-Experimental Methods	Pre-Post	Measure how program participants improved (or changed) over time.	Program participants themselves—before participating in the program.	The program was the only factor influencing any changes in the measured outcome over time.	Before and after data for program participants.
	Simple Difference	Measure difference between program participants and non-participants after the program is completed.	Individuals who didn't participate in the program (for any reason), but for whom data were collected after the program.	Non-participants are identical to participants except for program participation, and were equally likely to enter program before it started.	After data for program participants and non-participants.
	Differences in Differences	Measure improvement (change) over time of program participants <i>relative</i> to the improvement (change) of non-participants.	Individuals who didn't participate in the program (for any reason), but for whom data were collected both before and after the program.	If the program didn't exist, the two groups would have had identical trajectories over this period.	Before and after data for both participants and non-participants.
	Multivariate regression	Individuals who received treatment are compared with those who did not, and other factors that might explain differences in the outcomes are "controlled" for.	Individuals who didn't participate in the program (for any reason), but for whom data were collected both before and after the program. In this case data is not composed of just indicators of outcomes, but other "explanatory" variables as well.	The factors that were <i>excluded</i> (because they are unobservable and/or have been not been measured) do not bias results because they are either uncorrelated with the outcome <i>or</i> do not differ between participants and non-participants.	Outcomes as well as "control variables" for both participants and non-participants.
	Statistical Matching	Individuals in control group are compared to similar individuals in experimental group.	<u>Exact matching</u> : For each participant, at least one non-participant who is identical on <i>selected characteristics</i> . <u>Propensity score matching</u> : non-participants who have a mix of characteristics which predict that they would be as likely to participate as participants.	The factors that were <i>excluded</i> (because they are unobservable and/or have been not been measured) do not bias results because they are either uncorrelated with the outcome <i>or</i> do not differ between participants and non-participants.	Outcomes as well as "variables for matching" for both participants and non-participants.
	Regression Discontinuity Design	Individuals are ranked based on specific, measurable criteria. There is some cutoff that determines whether an individual is eligible to participate. Participants are then compared to non-participants and the eligibility criterion is controlled for.	Individuals who are close to the cutoff, but fall on the "wrong" side of that cutoff, and therefore do not get the program.	After controlling for the criteria (and other measures of choice), the remaining differences between individuals directly below and directly above the cut-off score are not statistically significant and will not bias the results. A necessary but sufficient requirement for this to hold is that the cut-off criteria are strictly adhered to.	Outcomes as well as measures on criteria (and any other controls).
	Instrumental Variables	Participation can be predicted by an incidental (almost random) factor, or "instrumental" variable, that is uncorrelated with the outcome, other than the fact that it predicts participation (and participation affects the outcome).	Individuals who, because of this close to random factor, are predicted not to participate and (possibly as a result) did not participate.	If it weren't for the instrumental variable's ability to predict participation, this "instrument" would otherwise have no effect on or be uncorrelated with the outcome.	Outcomes, the "instrument," and other control variables.
Experimental Method	Randomized Evaluation	Experimental method for measuring a causal relationship between two variables.	Participants are randomly assigned to the control groups.	Randomization "worked." That is, the two groups are statistically identical (on observed and unobserved factors).	Outcome data for control and experimental groups. Control variables can help absorb variance and improve "power".

Annex 4: Summary of Selected Studies Using “Objective” Indicators of Corruption

Source: adapted from Fazekas, Tóth, & King, 2013 and Fox 2014

Study	Indicator used	Country	Sector	Findings
Public Procurement and Rent-Seeking: The Case of Paraguay (Auriol et al., 2011)	Exceptional procedure type	Paraguay	general procurement	Entrepreneurs' economic incentives are distorted toward unproductive activities as the result of favoritism in the allocation of public contracts.
Active and Passive Waste in Government Spending: Evidence from a Policy Experiment (Bandiera, Prat, & Valletti, 2009)	Price differentials for standard goods purchased locally or through a national procurement agency	Italy	various standardized goods (e.g. paper)	Some public bodies pay systematically more than others for observationally equivalent goods and such price differences are sizeable. Differences are correlated with governance structure: the central administration pays at least 22% more than semi-autonomous agencies.
Building Political Collusion: Evidence from Procurement Auctions (Coviello & Gagliarducci, 2010)	Number of bidders Same firm awarded contracts recurrently Level of competition	Italy	general procurement	One extra term in office decreases the number of bidders and the winning rebate. The probability that the same firm is awarded more auctions, or that the winning firm is local, increases with time in office.
The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires (Di Tella & Schargrodsky, 2003)	Difference in prices of standardized products such as ethyl alcohol	Argentina	healthcare	Prices paid by hospitals for basic, homogeneous inputs decrease by 15 percent during the first 9 months of a corruption crackdown. After this period prices increase, but they are still 10 % lower than those prevailing before the crackdown.
Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. (Ferraz & Finan, 2008)	Corruption uncovered by federal audits of local government finances	Brazil	federal-local transfers	Electoral accountability – politicians were not re-elected if the public was made aware of their corrupt activities.
Proposal for a New Measure of Corruption, illustrated with Italian data (Golden & Picci, 2005)	Ratio of physical stock of infrastructure to cumulative spending on infrastructure	Italy	infrastructure	Where the difference is larger between the monies spent and the existing physical infrastructure, more money is being siphoned off to mismanagement, fraud, bribes, kickbacks, and embezzlement; that is, corruption is greater.
Politically Connected Boards of Directors and The Allocation of Procurement Contracts (Goldman et al., 2013)	Political officeholders' position on company boards	USA	general procurement	Companies that are connected to the winning (losing) party in national elections are significantly more likely to experience an increase (decrease) in procurement contracts.
Politics and Procurement: Evidence from Cleaning Contracts (Hyytinen et al., 2008)	Number and type of invited firms Use of restricted procedure	Sweden	cleaning services	Left-wing majority councils seem to invite fewer firms in restricted-entry auctions. Left-wing majority councils are also 1.5 as price sensitive as right-wing councils.
Corruption and the costs of redistribution: Micro evidence from Indonesia (Olken, 2006)	Difference between the quantity of in-kind benefits (rice) received, according to official records and reported survey evidence	Indonesia	welfare spending	18 percent of the rice distributed in a welfare program appears to have disappeared. Ethnically heterogeneous and sparsely populated areas are more likely to be missing rice.
Monitoring Corruption: Evidence from a Field Experiment in Indonesia (Olken, 2007)	Differences between the officially reported and independently audited prices and quantities of road construction projects	Indonesia	infrastructure (roads)	The central audit works through community response (social sanctions and village elections). Top-down and bottom-up accountability are synergistic.
Local Capture: Evidence From a Central Government Transfer Program in Uganda. (Reinikka & Svensson, 2004)	Difference between block grants received by schools according to official records and user survey	Uganda	education	Less leakage in block grants received by schools once missing funds were announced by local media.

Annex 5: Institutional Transparency Indicators

Source: Author

Fiscal transparency/ Budget	<p>Comprehensiveness of information included in budget documentation (PEFA PI-6) Transparency of inter-governmental fiscal relations (PEFA PI-8) Public access to key fiscal information (PEFA PI-10) Quality and timeliness of in-year budget reports (PEFA PI-24) Quality and timeliness of annual financial statements (PEFA PI-25) Timely, comprehensive, and periodic publication of budget documents (Open Budget Disaggregated Indicators) Publication of key budget documents (Open Budget Index):</p> <ul style="list-style-type: none"> • Pre-Budget Statement • Executive's Budget Proposal • Enacted Budget • Citizens Budget • In-Year Reports • Mid-Year Review • Year-End Report • Audit Report <p>% of community who perceive budget process to be transparent/information to be easily accessible.</p>
Procurement	<p>% of contract awards published [in a timely manner, by type of procurement/threshold value] % of procurement plans published Degree of access to information (MAPS 11) Transparency, competition and complaints mechanisms in procurement (PEFA PI-19) % of e-procurement log-ins that results in download of information package % of business community who perceive the public procurement system to be transparent</p>
Tax/Revenue	<p>Publication of activities, results and plans (TADAT P9.27) [Reporting on (i) financial performance, (ii) operational performance; Extent of publication of Tax Administrations future plans] Transparency of taxpayer obligations and liabilities (PEFA PI-13) % of clients/users who perceive tax administration to be transparent</p>
Public management and employment	<p>Internal orienting documents (strategy, guidelines, code of conduct, etc.) are easily accessible to all staff HR policies on recruitment, performance evaluation, and salary levels are clearly written and accessible to all staff</p>

Annex 6: Institutional Accountability Indicators

Source: Author

Budget	<ul style="list-style-type: none"> % of budgetary institutions preparing standardized internal audit reports % of accounts and audits backlog reduced Extent of unreported government operations (PEFA PI-7) Oversight of aggregate fiscal risk from other public-sector entities (PEFA PI-9) Effectiveness of internal audit (PEFA PI-21) Availability of information on resources received by service-delivery units (PEFA PI-23) Scope, nature and follow-up of external audit (PEFA PI-26) Legislative scrutiny of the annual budget law (PEFA PI-27) Legislative scrutiny of external audit reports (PEFA PI-28)
Procurement	<ul style="list-style-type: none"> % of legitimate complaints % of business community who think that there is an equal opportunity to receive public contracts # of complaints filed as a share of contracts awarded The country has systems and procedures for collecting and monitoring national procurement statistics (MAPS 5b) Quality control standards are disseminated and used to evaluate staff performance and address capacity development issues (MAPS 5d) Existence of contract administration and dispute resolution provisions (MAPS 8) Effective control and audit systems (MAPS 9) Efficiency of appeals mechanism (MAPS 10) Ethics and anti-corruption measures in place (MAPS 12)
Tax/Revenue	<ul style="list-style-type: none"> Identification, assessment, ranking & quantification of compliance risks (TADAT P2.3) Identification, assessment and ranking of institutional risks (TADAT P2.4) Designing, implementing, monitoring & evaluating risk mitigation activities (TADAT P2.5) External oversight of the Tax Administration (TADAT P9.24) Level of internal controls (TADAT P9.25) Public perception of integrity (TADAT P9.26)
Public management and employment	<ul style="list-style-type: none"> Conflict-of-interest restrictions exist for civil service, in code of conduct, civil service regulations, or separate legislation (PAM COI) % of agencies that have a functioning office/officer that provides guidance to public officials on how to avoid and/or mitigate conflicts of interest % of public officials trained on issues of conflict of interest, ethics, and internal anti-corruption measures % of public officials required to file financial disclosures who actually do so # of disciplinary actions initiated for violations of code of conduct



*Empowered lives.
Resilient nations.*

UNDP partners with people at all levels of society to help build nations that can withstand crisis, and drive and sustain the kind of growth that improves the quality of life for everyone. On the ground in more than 170 countries and territories, we offer global perspective and local insight to help empower lives and build resilient nations.

Copyright © 2015
United Nations Development Programme One United Nations Plaza
New York, NY 10017, USA

Graphic designer: Jacqueline Broner

For further information please contact:
Anga R. Timilsina (Ph.D.)

Programme Manager
UNDP's Global Anti-corruption Initiative (GAIN)
anga.timilsina@undp.org