UNDP

Empowered lives.
Resilient nations.

WHY, WHAT AND HOW TO MEASURE?
A USER'S GUIDE TO MEASURING RULE OF LAW, JUSTICE AND SECURITY PROGRAMMES

FOREWORD

In September 2012, Member States in the General Assembly of the United Nations reaffirmed that the rule of law is critical for sustainable development. Indeed, the rule of law is an important factor in accelerating achievement of the MDGs and will be essential to the post-2015 Development Agenda as both an enabler and an outcome of development in its own right. It is now beyond question that improving safety for individuals and communities, and providing access to fair and well-functioning legal systems that adhere to international human rights standards, are necessary to promote economic investment, prevent violence and conflict, encourage inclusive growth and eradicate poverty. To these ends, governments, civil society groups and multilateral actors have in many cases increased the resources specifically devoted to the rule of law area.

Within the UN system, the Secretary-General appointed DPKO and UNDP as the Global Focal Point for Police, Justice and Corrections in post-conflict and other crisis situations in order to harness rule of law expertise and deliver assistance under a coordinated UN framework. However, despite collective international attention and diverse development efforts in this area, data are often not collected or contextualized to measure effectiveness and contribute to evidence-based policy and programming. As a result, changes in the rule of law have not been easy to capture.

The rule of law is multi-dimensional and has numerous functions - so much so that it is often difficult to measure with time-bound indicators. Differences also exist when determining the achievement of programme outputs versus the impact of programmes on beneficiaries, or the overall attainment of rule of law at country and local levels. These challenges have contributed to a deficit in programme and project measurement. This Guide addresses these difficulties and provides practical direction for the use of data for evidence-based programming and for results reporting against established baselines, which aligns with the increased emphasis on data and measurement in UNDP's Strategic Plan (2014-2017).

Taking the need for measurement as a starting point, this Guide provides the answers to a wide range of questions: What are the needs for rule of law assistance? What type of programme should be designed? Has the programme been successful in achieving its stated goals and objectives? What is the impact of the rule of law programme for the beneficiaries? What impact has the programme had on the rule of law system in its entirety? Answering these questions requires both capacity for data collection and the willingness to do so, even if it means facing unfavourable results. Accordingly, this Guide works from the conviction that results-oriented development cannot be implemented without measurement. It is impossible to know if a programme triggered any change without collecting the necessary baseline data.

The Guide provides a wide range of suggestions for conducting measurement in data-poor and politically challenging environments. Complementing the UN Rule of Law Indicators, it makes a compelling case for the benefits of measurement, and also provides a realistic overview of the requirements for effective programme measurement. It provides valuable advice and numerous examples of how UNDP programme results can be obtained even when operating under budget, time and data constraints — all with the intent of encouraging greater attention to data collection by governments, UNDP, the UN Global Focal Point mechanism and other development partners.

It is our hope that this Guide will contribute to more tangible development outcomes resulting from more rigorous data collection, context-driven policy development, programme design and implementation.

Jordan Ryan, BCPR   **Magdy Martinez-Soliman, BDP**

TABLE OF CONTENTS

A2J	Access to Justice
BCPR	Bureau of Crisis Prevention and Recovery
BDP	Bureau of Development Policy
DDR	Disarmament, Demobilization and Reintegration
DIE	Deutsches Institut für Entwicklungspolitik (German Development Institute)
DPKO	Department for Peace Keeping Operations
EBRD	European Bank for Reconstruction and Development
ICA	Institutional and Context Analysis
IDB	The Inter-American Development Bank
IMF	International Monetary Fund
MDGs	Millennium Development Goals
NeCGD	Non-equivalent Control Group Design
OHCHR	Office of the High Commissioner for Human Rights
PAR	Participatory Action Research
RDD	Regression Discontinuity Design
RoL	Rule of Law
RoIJS	Rule of Law, Justice and Security Unit (UNDP/BCPR)
QualMM	Qualitative Measurement Methods
QuantMM	Quantitative Measurement Methods
SG	Secretary General
SICA	Central American Integration System
SGBV	Sexual and Gender-Based Violence
UNDG	United Nations Development Group
UNICEF	United Nations Children's Fund
UNODC	United Nations Office on Drugs and Crime
USAID	United States Agency for International Development

Authors:
Besiki Kutateladze
Jim Parsons
Vera Institute of Justice

INTRODUCTION

INTRODUCTION

For the United Nations, the rule of law refers to a principle of governance in which all persons, institutions and entities, public and private, including the State itself, are accountable to laws that are publicly promulgated, equally enforced and independently adjudicated, and which are consistent with international human rights norms and standards. It requires, as well, measures to ensure adherence to the principles of supremacy of law, equality before the law, accountability to the law, fairness in the application of the law, separation of powers, participation in decision-making, legal certainty, avoidance of arbitrariness and procedural and legal transparency.

UN Secretary-General, 2004

For UNDP, an agency driving sustainable development that helps countries eradicate poverty and significantly reduce inequalities and exclusion, rule of law (RoL) is both a means of achieving sustainable human development and the desired end of many UNDP programmes. As such, RoL is at once a development tool and a key enabler of the MDGs[1] and the successor development framework post-2015. UNDP supports rule of law, justice and security programmes — including the legal empowerment of the poor — in more than 100 countries worldwide. This work is central to achieving UNDP's Strategic Plan (2014-2017), which seeks to ensure that citizen expectations for voice, development, the rule of law and accountability are met by stronger systems of democratic governance and that countries have strengthened institutions to progressively deliver universal access to basic services. These objectives require that UNDP supports countries to ensure that legal frameworks prohibit discrimination as well as build the capacity of rule of law institutions in order to improve access to justice and redress. It further demands that interventions help empower communities and enable security services to increase citizen safety and reduce levels of armed violence. Supporting increased progress in gender equality and women's human rights is closely linked to the rule of law and a cross-cutting objective of UNDP's efforts. Strengthening the rule of law in such a manner is critical in crisis affected countries as a key element of state and peacebuilding.[2]

While the importance of RoL is widely accepted, determining whether reforms are effective and impactful on the country level is often not systematically analysed. Part of the reason for this lack of clarity is the commonly-held view that RoL is a difficult field to measure. However, since RoL is an essential building block for the development, prosperity and well-being of individuals, communities and states across the globe, the importance of reliable information on the impact of RoL initiatives far outweighs the challenges of measurement. An investment in the measurement of RoL initiatives promises a greater return for their beneficiaries while limiting the potential for unintended, negative consequences that can come with misdirected interventions.

---

1   These are eight international development goals that the UN agreed to achieve by 2015. The goals are: (1) eradicating extreme poverty and hunger; (2) achieving universal primary education; (3) promoting gender equality and empowering women; (4) reducing child mortality rates; (5) improving maternal health; (6) combating HIV/AIDS, malaria, and other diseases; (7) ensuring environmental sustainability; and (8) developing a global partnership for development.

2   See Changing with the World: UNDP Strategic Plan 2014-2017 and its Integrated Results and Resources Framework, in particular Outcomes 2 and 3 and related outputs 2.6, 3.4 and 3.5, as well as Outcome 4, output 4.2. http://www.undp.org/content/dam/undp/library/corporate/UNDP_strategic-plan_14-17_v9_web.pdf Report of the Secretary-General on the Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies (S/2004/616), pg. 4.

The terms rule of law, justice and security are widely used and are often overlapping. This Guide is based upon the Secretary General's definitions of the "rule of law",[3] "justice"[4] and "security".[5] RoL is used throughout this publication to refer to the framework of support the UN provides at national level, which includes all three programming areas.

What is the Guide For?

This Guide is intended to improve the effectiveness of RoL programming with a view to implementing UNDP's Strategic Plan. It furthers the commitment of the organization in the Plan to ensure the highest standards of delivery of development results through rigorous monitoring and quality assurance and measurement. The Guide also responds to an increasing demand from a wide range of national stakeholders and UNDP Country Offices for guidance on how to measure the impact of RoL programmes. This is the first guide that focuses specifically on the measurement of RoL programmes and projects across the spectrum of development settings, including conflict-affected and fragile environments. In many cases, UNDP programme staff will not be collecting and analysing data themselves, but hiring external consultants to do so. Guidelines for measuring the impact of RoL programming will help RoL practitioners and consultants determine, for example: how to operate under budget, time, political and data constraints; what types of skills to look for when hiring external expertise; how to use research findings to design and implement effective programmes; and how to translate measurement findings into practice. Ultimately, the strategies laid out in this Guide will contribute to building a culture of accountability and transparency within and between UNDP, donors, national governments and civil society partners.

> This Guide provides practical guidance to UNDP Country Office staff and national stakeholders on ways to measure the effectiveness of rule of law programmes. It details the benefits of rigorous measurement and provides an overview of data collection methodologies, their feasibility, limitations, and practical considerations.

---

3   Report of the Secretary-General on the Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies (S/2004/616), pg. 4.

---

4   "For the United Nations, 'justice' is an ideal of accountability and fairness in the protection and vindication of rights and the prevention and punishment of wrongs. Justice implies regard for the rights of the accused, for the interest of victims and for the well-being of society at large." Ibid.

---

5   Security sector "is a broad term often used to describe the structures, institutions and personnel responsible for the management, provision and oversight of security in a country. It is generally accepted that the security sector includes defense, law enforcement, corrections, intelligence services and institutions responsible for border management, customs and civil emergencies. Elements of the judicial sector responsible for the adjudication of cases of alleged criminal conduct and misuse of force are, in many instances, also included. Furthermore, the security sector includes actors that play a role in managing and overseeing the design and implementation of security, such as ministries, legislative bodies and civil society groups. Other non-state actors that could be considered part of the security sector include customary or informal authorities and private security services." Report of the Secretary-General on SSR, entitled 'Securing Peace and Development: The Role of the United Nations in Supporting Security Sector Reform', dated 23 January 2008 (A/62/659/-S/2008/39).

What is the Scope of the Guide?

While the Guide offers practical guidance on a wide range of measurement topics, it is designed to be used in combination with other resources (see Appendix G). The Guide does not provide a comprehensive description of existing data collection activities, indicator initiatives or other measurement tools. This Guide does provide examples of measurement tools and research designs, with suggestions on how existing approaches can be tailored to suit the needs of rule of law practitioners. The Guide reviews commonly used methodologies that can be adapted to data collection and analyses in fragile and post-conflict environments. It offers suggestions for ways to overcome a wide range of practical challenges to measurement, including budget and time constraints, problems with the availability and quality of data, and a lack of stakeholder support.

This Guide is not meant to be a textbook. It is a practical tool addressing real world challenges faced by people working in the development field. Wherever possible, the Guide incorporates case studies and examples drawn from UNDP's work.

Who is the Guide For?

The primary audience for this Guide is RoL practitioners based in UNDP country offices, regional centres and headquarters.[6] However, the impact of improved measurement outlined in the Guide should be primarily for the benefit of national stakeholders.

This Guide will be useful to national stakeholders involved in various RoL reform activities, including government agencies, civil society and academic institutions. Forging partnerships to improve measurement will help ensure that data collection activities are based on local experience and reflect national priorities while enhancing local ownership of a project that aims to (re)build local capacity for research and innovation.

This Guide will also be useful to donor agency staff. Funders and development partners in general increasingly require more robust performance measures that document a social return on their investments. However, these agencies also struggle to measure governance programming, including for RoL initiatives and especially in post-conflict and fragile environments. Although many funders are developing their own tools for measuring efficacy, this Guide is intended to contribute to greater programme effectiveness and accountability.

---

6   The audience of this Guide is the same as for the Handbook on Planning, Monitoring and Evaluating for Development Results, (UNDP, 2009).

How is the Guide Organized?

The Guide consists of three chapters and an appendix. The first chapter is entitled 'Why Measure'. It details the benefits of measurement, including the importance of data- and results-driven programming, the need for baseline data collection to track programme impact as well as the usefulness of evaluation tools to expand successful RoL initiatives. The second chapter is entitled 'What to Measure'. It provides practical guidance on how to determine the purpose and scope of measuring the success of a development programme. This chapter also provides direction on engaging stakeholders, and assessing the feasibility of a research design. The third chapter is entitled 'How to Measure'. It describes common measurement and evaluation approaches, provides guidance on data collection and analysis, and offers suggestions for translating findings into policy recommendations. Finally, the appendix

FIGURE 1: THE GUIDE IN BRIEF
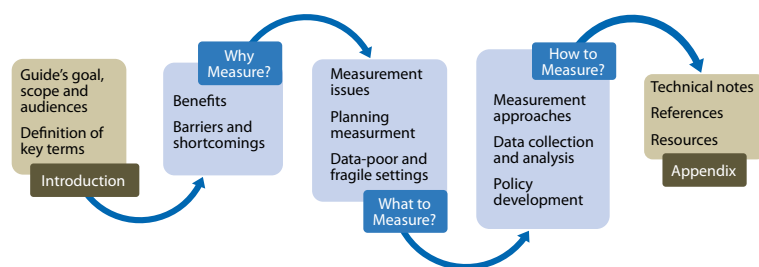
lists existing measurement tools and other resources.

This Guide is divided into five standalone sections that need not be read sequentially. Those with limited time can review the sections that are most relevant to their needs. However, the following section outlines several fundamental terms necessary for understanding the entirety of this document.

What is 'Measurement'?

Research terms, such as measurement, monitoring, assessment, survey and evaluation, are commonly, but inconsistently, used in the development field to refer to a range of activities. In this section, key terms used throughout this document are defined in accordance with UNDP's corporate guidance on monitoring and evaluation. This Guide further builds on these terms by specifying two stages of evaluation discussed below. These definitions serve the practical purpose of distinguishing among different types of measurements used in development programming and applied to the RoL area.

In this Guide, we use **'Measurement'** in the context of three different steps: **Assessment**, **Mid-Term Evaluation** and **Final Evaluation**.[7] These three steps to measurement are based on a wide range of activities that include data collection, analyses and interpretation of findings. These are complementary steps that exist as a part of a continuum; one is not an alternative to another. For instance, baseline data that is collected during the assessment phase should be compared with mid-term and final evaluation data, as shown below. Although practical constraints often limit data collection, effective programming should include elements of all three measurement steps.

FIGURE 1: THE GUIDE IN BRIEF

Assessment		Mid-term Evaluation		Final Evaluation
Baseline data	→	Mid-point data to compare with baseline data	→	Follow-up data to compare with baseline and mid-point data
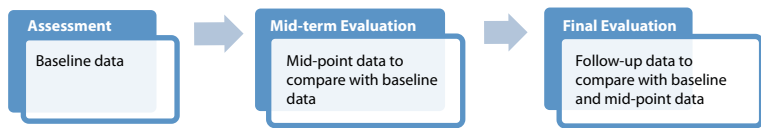
Defining 'Assessment'

In this Guide, **'Assessment'** refers to a set of data collection and analytical activities typically completed during the pre-implementation design phase of a project. This kind of assessment generates baseline data, essential for documenting change over time, whether it is positive or negative. Without such baseline data there is no way of knowing if a project resulted in a change. Assessments are also used to determine how to design a project to minimize negative, unintended consequences and achieve desired outcomes such as increasing court access for women or decreasing the number of people in pre-sentence detention. Assessments can also inform the design of ongoing evaluation activities by identifying programme objectives, assessing the availability of data, and designing mid-term and final evaluation measures. RoL development is a highly politicized field and assessments can also help pre-empt and plan for potential political obstacles to development initiatives. The assessment methods discussed in this Guide are a necessary complement to other generic UNDP assessment tools that relate more closely to gender, conflict, capacity and other specific needs (see Appendix F for additional information).

---

7   Currently, these terms are used in numerous ways. Therefore, this Guide also seeks to clarify them in relation to measurement. There are, of course, other types of measurements like the Institutional and Context Analysis Guidance Note (UNDP, 2012) as well as other measurements that look at risk or conflict. These measurements should be used in tandem with measurement approaches proposed in this Guide.

An assessment can serve the following four purposes:

Purpose 1	Collect baseline data to track progress over time
Purpose 2	Explore and document a problem
Purpose 3	Help determine what type of project to implement for a positive impact
Purpose 4	Inform the design of mid-term and final evaluations
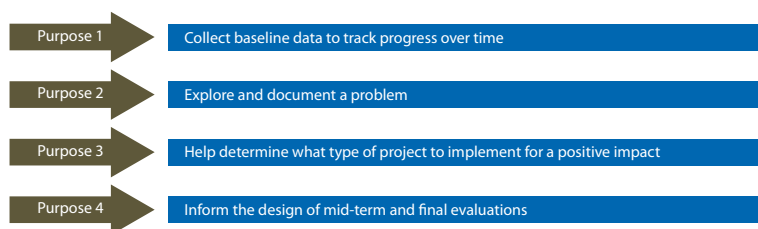
BOX 1: GUIDING PRINCIPLES FOR UN RULE OF LAW ASSISTANCE

Process: In 2010, UNDP Guinea Bissau carried out an access to justice assessment under the auspices of the State of Guinea-Bissau/Ministry of Justice. Its purpose was to gather and define baseline data for UNDP's RoL programme as well as to inform national policies on access to justice. The assessment included both quantitative and qualitative analyses and produced recommendations on improving access to justice in target areas. Fieldwork involved expert surveys with representatives of traditional justice authorities, government partners and NGOs/grassroots organizations. Additionally, a population survey was conducted and focus group discussions were held with vulnerable groups, including women, children, refugees, prisoners and the disabled. The assessment identified the obstacles and difficulties in accessing justice encountered by vulnerable groups and those responsible for the provision of justice services.

Results: Access to justice was found to be limited for the wider population, especially for women and children. Poor infrastructure, lack of legal awareness, an outdated legislative framework and lack of coordination between state and traditional justice mechanisms were all contributing factors. The assessment also found the effects of certain cultural beliefs and traditional practices to be a further obstacle to accessing the formal justice system as well as for safeguarding human rights.

Given these baseline findings the assessment was able to make a number of recommendations that included strengthening the state presence and establishing justice institutions in the regions. Ultimately, baseline data was essential for measuring the impact of Legal Aid Centres supported by the project. Recommendations were also taken into account by the Government for the development of their national justice strategy in 2011.

Defining 'Mid-term Evaluation'

UNDP defines evaluation as "a rigorous and independent assessment of either completed or ongoing activities to determine the extent to which they are achieving stated objectives and contributing to decision-making."[8] The use of the terms 'mid-term' and 'final' evaluations in this report is consistent with this definition.

A mid-term evaluation is used during the lifespan of a project to help determine if project activities are reaching milestones, have been consistent with the initial plan, and if they are still likely to result in desired outcomes, or require modification.[9] For a project using performance indicators, the mid-term evaluation will help track change over the course of the project and in comparison to baseline data. While the mid-term evaluation should be ongoing, the frequency of data collection will depend on available resources, the length of the project, or the nature of anticipated outcomes.

Mid-term evaluations can serve the following three purposes:

Purpose 1	Ensure project is making progress and meeting its milestones
Purpose 2	Determine if the project is being implemented according to plan
Purpose 3	Determine if and how to modify or terminate the project
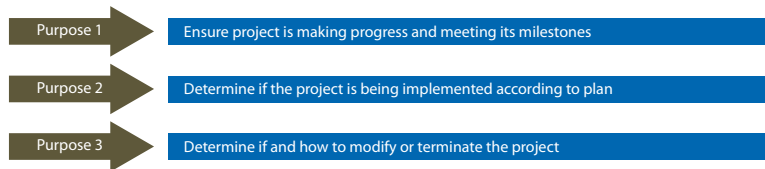
Mid-term evaluation has some similarities with 'Monitoring' as used within the UNDP system. Both mid-term evaluation and monitoring can serve the three purposes outlined above, involve some type of data collection and can be implemented throughout the lifespan of a project. However, unlike monitoring, mid-term evaluation uses more thorough data collection methods and often includes data that is designed to quantify programme impacts to provide a more rigorous measure of effectiveness. Therefore, findings generated as a result of mid-term evaluation tend to be more accurate and easier to generalize.

This Guide does not address Monitoring (see *Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009)* for monitoring steps). We suggest that mid-term evaluations and monitoring activities are developed in tandem to avoid any duplication of efforts.

8   See Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), p.8.

9   Often, when a project is designed, project staff may not have all the information necessary to determine the most effective project plan. Furthermore, changes in a political, economic or security environment may require the plan to be reconsidered. Therefore, as more information becomes available while implementing the project, changing certain activities may be necessary. In rare cases, project staff may have to discontinue the project entirely, especially if it becomes apparent that the project is not achieving its goals or is having a potentially detrimental impact.

Defining 'Final Evaluation'

Final evaluation will help project staff determine if activities led to, or are associated with, desired results. A final evaluation can be conducted at the end of a project, upon its completion, or a few months later, depending on whether the focus is in short-, medium- or long-term results. Regardless, planning for a final evaluation should occur during the assessment phase of any project to ensure the collection of necessary information.

A carefully designed final evaluation will reveal whether the project achieved its goals and what project and contextual elements predict success. For example, it may be discovered that a programme that sought to increase legal representation was more effective in larger urban courtrooms or with defendants facing civil charges. This information is valuable when considering how to improve a partially successful project or where to expand services to maximize impact. To determine whether a project was effective reviewers need pre-determined criteria for success (i.e., how much improvement is needed to count the programme as achieving its aims). These should be defined at the beginning of a project.

Final evaluations and mid-term evaluations are complementary, often using the same measures and data collection methods. They provide important information to maximize impact by providing an opportunity to correct implementation problems before they derail projects. Evaluations can be used to expand successful initiatives and communicate lessons learned, to help others avoid pitfalls and to overcome challenges faced by pioneering UNDP projects.

Below are three purposes that both mid-term and final evaluations serve:

Purpose 1	Determine if the project achieved its stated goals and planned results
Purpose 2	Understand why results were not achieved, or were achieved only partially
Purpose 3	Use findings, both positive and negative, to design effective projects
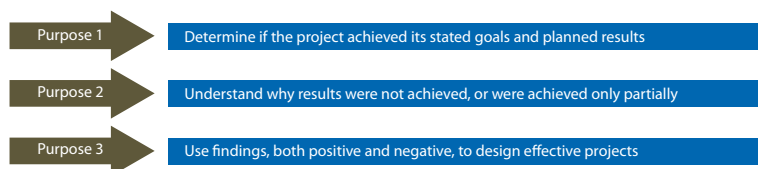
Throughout this Guide we will be referring to assessment, mid-term evaluation and final evaluation as different components of a 'measurement' strategy.

TABLE 1: PROJECT MEASUREMENT ELEMENTS			
Measurement	When to Initiate	Purpose	Activities and Process
Assessment	Before the project begins and preferably before it is even designed.	Collect baseline data; explore problems and define needs; inform project development.	Examine a problem, determine how to best address it, identify indicators, collect baseline data, identify desired outputs and outcomes, identify potential unintended consequences, decide how to structure a project, estimate required resources.
Mid-term Evaluation	Throughout project implementation.	Measure progress; decide to continue, modify, or terminate a project.	Determine if project activities have been consistent with the initial plan and if they are leading to the desired outcomes.
Final Evaluation	Can be close to project completion, upon project completion or a few months after completion.	Determine if a project resulted in desired outcomes.	Helps decide if all activities specified during the assessment phase were implemented (and if not why not); if the desired outcomes were fully or partially achieved; if the project had an impact or led to some change; and if unintended consequences were avoided or minimized.

WHY MEASURE?

CHAPTER 1

WHY MEASURE?

Designing and guiding to completion a RoL programme without any data is much like feeling one's way in the dark. Effective measurement is essential for good development practice and this is reflected in the increasing emphasis on the use of data to inform the design and evaluation of RoL programmes. The three phases of measurement — assessment, mid-term evaluation and final evaluation — will generate the type of data that can help design and conduct a successful project. Without information on the problem at hand, the state of affairs that preceded a project, the obstacles that may appear along the way, or the benefits that accrue to the people served by the project, it will be difficult to have any sense of whether the intervention was successful.

A well-designed measurement strategy will ensure that a programme is based on an objective assessment of the development challenges at hand and full knowledge of what is required to overcome obstacles and meet implementation goals. A clear and well-designed measurement sequence that begins with an assessment, moves to mid-term evaluation(s) and ends with a final evaluation can help to guarantee the effectiveness of a project and help ensure transparency and accountability.

However, establishing this three-tiered measurement strategy requires a significant investment of time and resources. It should, therefore, be clearly articulated why collecting information is important to ensure the participation and support of a range of constituents, including project beneficiaries and national stakeholders.

1.1. Improving Data Collecting Capacity

An added benefit of collecting and analysing data is that it encourages a commitment to measurement for UNDP country offices, international partner organizations, national governments and civil society.[10] For example, building a database to monitor the length of time that people are held in pre-trial detention, as part of a UNDP court improvement project, may help courts to better measure and track their performance on an ongoing basis. Moreover, UNDP can help to build local stakeholders' capacity for research, data collection and management. For example, funding local civil society organizations or academic institutions to conduct public surveys can eventually benefit UNDP projects by creating a pool of local experts and in-country resources to conduct similar projects in the future.

---

10   UNDP has provided support to strengthening case management capacity across the penal chain in countries including DRC, Guatemala, Haiti, Iraq, Liberia, Nepal, the occupied Palestinian territory, Sierra Leone, Somalia and Timor-Leste. See Global Programme Annual Report 2011 (UNDP, 2011), p.48.

In addition, requesting data from government agencies can trigger their internal data collection processes. For example, as a part of a pre-trial detention project data might be requested from local prisons on the percentage of pre-trial detainees visited by a lawyer within one month of their transfer to prison. Creating a demand for this type of information will increase the likelihood that the information will be collected and used on an ongoing basis. In many countries, it may be possible to support and supplement the work of national statistical offices or crime observatories as a capacity building measure.

1.2. Building Stakeholder Support and Ownership[11]

RoL projects should be owned, first and foremost, by national stakeholders, including governments, civil society organizations and members of the public. Project effectiveness depends on the support and ownership of local government agencies, civil society and community leaders. Without the support of key stakeholder groups, RoL initiatives may not be sustainable. Demonstrating that a project is effective can help secure their sense of ownership. Furthermore, participation by local stakeholders in the assessment phase at the beginning of a project may help to ensure their buy-in from the outset.

Moreover, in times of scarce resources, it is increasingly difficult to convince donors to fund specific projects. Measurement can inform funding decisions by directing support towards effective projects, closing down initiatives that fail to achieve their goals, improving the impact of investments and maximizing value for money. Increasingly, baseline assessment data is an essential component of funding applications.

1.3. Informing Project Design

To be effective, the design and targeting of projects should be based on data collected during the assessment phase of measurement. For instance, a project to establish mobile courts in remote and rural parts of a country could use information from the initial assessment to target areas where people lacked access to courts. A project that proceeds without an initial assessment of circumstances may look good on paper but falter on the ground during the implementation phase. Furthermore, unanticipated obstacles may only become apparent once a project is up and running. Collecting data on an ongoing basis also provides an important mechanism for fine-tuning project design. In the case of mobile courts, for example, their effectiveness may be limited by problems with accessing conflict-affected areas within a country, misunderstandings within communities about the types of cases that can be handled by mobile courts, or cultural barriers that prevent women from testifying

---

11   See 2.3.B. on how to identify and engage stakeholders.

in public settings where men are present. A well-designed measurement exercise will allow project teams to identify and adapt to such emergent obstacles. Rigorous assessment conducted at the outset of a project can ensure that any intervention reflects the needs of local stakeholders and beneficiaries, rather than the preferences of development agencies and other international organizations.

1.4. Gauging Project Effectiveness and Informing Practice

Collecting data is often the only way to know if a programme is strengthening the rule of law. By identifying successful programmes, measurement can be used to expand initiatives to serve more people and have a greater impact. Furthermore, by testing programmes and describing approaches that are effective in improving security and justice in challenging, conflict-affected environments, evaluations can contribute to future, evidence-based programming. This can avoid the common problem of using template activities that may not be suited to the local political or cultural context, or having to 'reinvent the wheel' when designing new RoL initiatives. For example, development agencies working in multiple countries recognize the need to improve linkages between state and non-state justice institutions as a way of increasing accountability and extending access to justice (A2J) to rural regions. However, there is currently a lack of research-based information on successful approaches to linking these systems. Once this research is done, however, it can inform the work of others developing similar projects in the same country and elsewhere. See Chapter Three for more approaches to measuring programme effectiveness.

1.5. Increasing Transparency and Accountability

Measurement can also help to make UNDP, its donors and subcontractors, national governments and non-governmental organizations more transparent and accountable. For example, by assessing the number of complaints that a police department responded to within a given time period, UNDP may help to create more effective and responsive police accountability mechanisms. In addition, by publishing Impact Evaluation reports that describe the impact of funding, UNDP not only enhances its own transparency, but can also model best practices for government agencies and other development actors who may need to publish spending reports. For instance, a state supported project on community security including efforts on small arms and light weapons control could set a precedent for accountability by publishing its total project budget along with the impact on security for the targeted communities, for example, through data on homicide rates as well as perception surveys in targeted communities. By using data to critically examine its work and providing information on both successful and unsuccessful projects, UNDP will increase its credibility with

national and international partners. Accountability is also directly linked to UNDP's ability to raise funds in the future: aid is public money and, as such, taxpayers in donor countries want to know if their investments yielded a positive social return.

1.6. Including the Perspective of Vulnerable and Marginalized Groups

Carefully structured assessments can help to incorporate the views and concerns of vulnerable and marginalized groups. In many countries, rural populations, women, young people and other groups are less visible in both a literal and political sense. As a result, projects that are developed and tested in urban areas and with government officials may fail to address their needs and concerns. For example, establishing re-porting desks for crimes against women in major police stations based in provincial capitals may be entirely ineffective for addressing gender-based violence in small rural communities where women have limited access to transportation. A thorough assessment of the needs of women living in small towns and villages could help to ensure that RoL programmes are attuned to their needs. Public surveys and focus groups are particularly effective methods of collecting data to illuminate the needs and opinions of diverse and under-served groups (see Section 3.2.B. and 3.2.C. for more on these methods).[12]

---

12 See *Programming for Justice: Access for All: A Practitioner's Guide to a Human Rights-based Approach to Access to Justice* (UNDP, 2005).

BOX 2: OVERCOMING COMMON CHALLENGES TO MEASUREMENT

While there are benefits to measuring RoL programmes, there are also challenges. Awareness of these challenges can lead to strategies to pre-empt and overcome them. Some common measurement challenges and suggested solutions include the following:

Cost challenges: Measurements are often costly and extensive ones can be very costly, requiring large amounts of data and thorough analyses. One common objection is that resources expended on measurement could instead be used to serve a larger group of project beneficiaries.

> Possible Solutions → Emphasize to donors and other key stakeholders that measurement provides the only way to ensure that RoL projects achieve their intended results. Measurement can also help to improve the overall quality and efficiency of a project by targeting those who will benefit the most and ensuring accountability for the money invested in programming. Ultimately, investment in a smaller set of effective RoL interventions achieves greater development results and value for public money than many projects that do not deliver real, sustainable change.

Challenges related to a lack of familiarity with measurement methods: Many people who work for local NGOs, national governments and international organizations may be unfamiliar with the steps necessary for measurement. In some cases, potential partners may try to block data collection because they feel excluded from the process of collecting and analysing data, or fear that their authority will be undermined if research data is used in the decision-making process. It is extremely important that assessments or data collection do not alienate governmental partners who are often keenly aware of serious shortcomings on the ground. These partners will often require reassurances to overcome their perception that measurement is a donor imposed requirement of limited value to them.

> Possible Solutions → To pre-empt negative perceptions it is crucial to manage assessment results very carefully. Including national stakeholders in measurement steps is often the best way to avoid accusations that programming is donor-driven, Western or otherwise imposed. Therefore, engage government officials early on in an assessment process to ensure that they understand the process and can trust the motivations behind it. Specifically, the following four steps can help to generate local support:
>
> (1) Include partners in initial discussions as a way of incorporating their interests and concerns into the design of measurement activities.
>
> (2) Brief senior officials and development partners about data collection plans and how the findings will be used.
>
> (3) Brief national stakeholders privately with any preliminary findings to encourage their ownership of the results.
>
> (4) Distribute findings and recommendations among all stakeholders in advance of their general release and provide an opportunity for the airing of comments and concerns.

Challenges of entrenched interests: In some cases, project staff may encounter obstacles to measurement because of a concern that the results will upset entrenched interests and disrupt established practices. Senior law enforcement officials, for example, may feel threatened if they feel that information on the number of complaints

BOX 2: OVERCOMING COMMON CHALLENGES TO MEASUREMENT (CONTINUED)

against the police will be used to gauge their performance.

> Possible Solutions → Structure the measurement exercise in a way that introduces incentives for positive change. For instance, draw attention to improved indicators such as a reduction in the number of police complaints. In some cases, however, it may be necessary to confront entrenched interests and seek support from other sources such as the media or civil society organizations.[13]

Challenges of maintaining good relations with local partners: Negative findings of a project's measurement may damage relationships with stakeholders who were also involved in the project design and implementation. For example, a project seeking to make the court system more sensitive to sexual and gender-based violence (SGBV) could highlight problems that local project partners may be unwilling to discuss. If a government partner finds out that a project they supported was unsuccessful this may tarnish the project's reputation and hinder the implementation of similar projects in the future.

> Possible Solutions → Involve local partners in every step of a project's implementation and measurement process so that they are apprised of preliminary findings and may offer their input. Try to avoid unnecessarily criticizing local partners and instead discuss findings in terms of what can proactively be done to achieve desired outcomes.

Challenges presented by the pressure to succeed: Pressure to demonstrate success could directly or indirectly influence the selection of a RoL project. This is a dangerous because there are areas where it is difficult to demonstrate clear measurable results, often because short timeframes do not allow changes to mature sufficiently to show tangible results. The fact that results will be difficult to measure should not provide a disincentive to pursue good development programming.

> Possible Solutions → Work with national stakeholders to develop programme areas before determining how to evaluate projects. Assessments can help to determine what projects should be priorities and how to implement them to achieve desirable changes. The ability to evaluate a project's effectiveness is an important consideration, but it should not dictate priority areas.

KEY TAKEAWAYS

Chapter One — Why Measure

• Use measurement to maximize impact; build data capacity; increase the transparency and accountability of UNDP and government agencies; generate stakeholder support and national buy-in; and capture perspectives of vulnerable and marginalized groups.
• Overcome potential obstacles to effective measurement by anticipating costs, preserving important relations when faced with unfavourable results, and resisting pressure to demonstrate success unduly.

---

13   For further information on understanding entrenched interests see *Institutional and Context Analysis Guidance Note* (UNDP, 2012).

WHAT TO MEASURE?

CHAPTER 2

WHAT TO MEASURE?

The first step in measurement is to identify key indicators and determine if they can even be measured. It is important to conduct baseline assessments to outline the services provided as well as outcomes sought, and to ensure that measures taken reflect the concerns of stakeholders and beneficiaries. This chapter lays out a process for conducting assessments, mid-term evaluations and final evaluations.

SECTION 1: KEY CONCEPTS AND ISSUES IN MEASUREMENT

2.1.A. Key Measurement Concepts and Rule of Law Questions

UNDP's RoL projects operate in diverse settings and address a wide variety of problems. While there is considerable variation in the scope, goals and budgets of these projects, each one includes a set of activities with a specific outcome in mind.[14] A well-designed measurement system can help to determine if the activities were appropriate, adequately implemented and impactful. A clear understanding of the difference between the theoretical potential of a project and the realities of implementation is vital to assessing whether projects work in practice.

Assessment, Mid-term Evaluation and **Final Evaluation** (defined earlier in 'Introduction — What is 'Measurement'?') can use various research designs (quantitative or qualitative[15]) and data sources (administrative data, public surveys, focus groups, etc.[16]). As with any research project, each of these measurement stages requires the use of research questions, a data collection methodology and an analysis strategy.

Assessments typically occur at the planning stage to collect baseline data and inform project design by assessing the scale of the problem, including the areas and populations that are most affected. It is often important to identify the needs of the project's target group in order to design and target services appropriately. Assessments can also highlight opportunities for a project, by identifying venues for reach-

---

14   See Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), Table 5 on p.53 for explanations of these terms and how they relate to activities on one end and impacts on the other.

15   For more on the differences, strengths and weaknesses of each measurement methodology, see: 3.1. Measurement Approaches.

16   For types of data and data collection tips, see: 3.2.A. through 3.2.F.

ing the target population, or locations that are likely to deliver the greatest impact. Also, assessments can identify potential barriers that may obstruct the delivery of services. And, perhaps most importantly, assessments can provide data with which to inform the project design, especially in terms of setting realistic baselines, targets and indicators of progress.

If using performance indicators (see section 3.1.A. for more) to measure project outcomes, they should be developed and baseline data should be collected before launching the project's activities. Indicators are particularly powerful if they are developed in partnership with local stakeholders who will then feel vested in the project and view the indicators as valid, credible and useful.

One of the first steps in the assessment process is to investigate existing data sources. It is only necessary to collect information if there are no relevant sources of data that meet the project's needs. For example, DPKO and OHCHR recently developed a series of indicators as a part of the UN Rule of Law Indicators Project that cover police, courts, prosecution, criminal defence and prisons. Other international projects, such as the regional Barometers (Afrobarometer, Arab Barometer, Asian Barometer,

---

BOX 3: QUESTIONS THAT SHOULD INFORM THE WAY THE ASSESSMENT IS DESIGNED

- Do government agencies collect relevant administrative data for use as baselines?
- Do civil society and other organizations already carry out public perception surveys?
- Are there national development goals with targets and indicators already in place?
- What is the scale of the problem that the proposed project will address?
- Is the problem widespread, or does it disproportionately impact particular regions or populations?
- What resources and capacities are required to implement the project and ensure sustained outcomes?
- What relationships are necessary for implementation?[17]
- Are there social norms such as those concerning gender, or political sensitivities that may present obstacles to measurement?
- Have evaluations of similar projects by UNDP been undertaken in the past? What were the lessons learned?
- Are other agencies designing or implementing similar projects?
- What information is needed to conduct an evaluation, and is this information currently available or will it require additional data collection?
- Can the project build upon existing data collection activities to minimize costs and avoid the duplication of data requests and collection?

---

17   On this point and the following ideally refer to an institutional and context analysis assessing the RoL context, or consider undertaking one in order to fully explore complex relationships, sensitivities and vested interests at play. See: Institutional and Context Analysis Guidance Note (UNDP, 2012).

Eurobarometer and Latinobaromoeter) conduct public surveys in multiple countries that cover a range of governance issues. Similarly, national governments, civil society groups and other international development agencies may conduct national censuses or public surveys, collect administrative records, or compile other types of data.

The next point in the measurement process is the Mid-term Evaluation, which includes measurement activities conducted during the life of a project. Box 4 below provides examples of the types of questions that can inform an assessment of project activities and show if a project is progressing towards its goals. Such an evaluation can provide critical early feedback about problems with implementation and can indicate whether existing activities need to be augmented or modified to ensure their effectiveness. While mid-term evaluations will not provide feedback on all desired results, they can provide early signs of project effectiveness.

BOX 4: TYPICAL MID-TERM EVALUATION QUESTIONS[18]

- Is the project reaching its targets? Do indicators show positive change from baseline data?
- How is it known that the project is having an impact?
- Is the project implemented as planned and what are the challenges for project implementation?
- What obstacles did the project encounter?
- Are modifications to the original plan required to overcome these obstacles?
- Are important stakeholders supportive and constantly engaged?
- Are there new stakeholders who should be engaged?
- Can the project be completed on time?
- What is the cost to date and what resources are required to complete implementation?
- What is the return on investment?
- What steps are required to replicate the project in other jurisdictions?
- What important relationships are required to sustain the project in the longer term?
- Does the project take into account the political situation, such as upcoming elections or demonstrations?[19]
- Did the project reach all members of the target population, including those living in remote or difficult to access locations?

18   Some of these questions are about activities while others are about outcomes. When conducting evaluations, capturing both is important for a fuller understanding of the programme implementation and effectiveness. For more on the results-based management results chain, see Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), Figure 9 on p.55. The figure shows how inputs, activities, outputs, outcomes and impacts connect with one another.

19   For more on looking more closely at political will and stakeholder incentives in this field, see: Institutional and Context Analysis Guidance Note (UNDP, 2012)

UNDP projects should be based on a clear understanding of the relationship between project activities, outputs, outcomes and impacts. Perhaps the most common question asked in measurement is whether a project was effective, i.e., did it achieve its intended outputs and outcomes? For example, did establishing a legal aid mechanism for SGBV survivors result in more survivors reporting their victimization and greater satisfaction with case outcomes? Similarly, can it be determined whether supporting the adoption of domestic violence law resulted in actual protection of women's rights? These are examples that require Final Evaluation, which typically assess the changes — positive or negative — compared to baseline assessments and mid-term evaluations. The timing of final evaluation activities should be informed by the nature of the project. For example, it may take some time for local residents to learn about improvements in accessibility resulting from a project to rehabilitate courthouses in rural areas.

BOX 5: TYPICAL FINAL EVALUATION QUESTIONS

- Did the project meet its objectives?
- Did the project benefit its clients?
- Are some people more affected than others?
- Are there any negative or unanticipated side effects?
- Is the benefit reasonable considering the project cost?
- Should the project be expanded?
- Could the project be sustained by national stakeholders?

When developing measures for RoL projects, it is important to consider the use of existing data sources. For example, it may be easy to access accurate records describing the number of female police officers, prisons, convictions and/or police arrests. Official records detailing the gender of serving police officers, for example, may be used to gauge whether UNDP's support for the recruitment of female police officers improved gender parity in the police force. This information could be used to calculate the percentage of officers that are female, and then compare the same figure with previous years. However, in many settings where basic information is not consistently or accurately maintained, it may be necessary to collect data independently and/or use multiple measures to control for weaknesses in the data (see Chapter 3 for more on data collection techniques and the use of multiple measures.)

Time lags between project activities and the expected impacts may lead to limited measurability. For example, the public safety benefits of training police investigators may take several months or years to emerge. In situations where the overall goal is not easy to measure given restrictions on time or resources, it is often possible to measure less ambitious outcomes or simply focus on outputs — for example, whether there was a measurable difference in case processing time in the six months following a police training programme.

2.1.B. Using Proxies and Multiple Measures of RoL Concepts

In some cases, when it is not possible to directly measure an impact or change, it will be necessary to develop proxy measures. Proxy measures act as a substitute when it is not possible to measure the desired outcome directly. For example, if it is difficult to measure increased confidence in the police, proxy measures — which may be associated with confidence — could include changes in the number of calls for police assistance, the number of witnesses volunteering to provide testimony, or public surveys measuring the perception of police trustworthiness. To measure an issue as complex as confidence, it is best to combine evidence from different data sources and proxies to provide multiple measures of the underlying concept. This approach is known as 'triangulation'.

For example, to understand the impact of mobile courts a project could:

- Interview a selection of community members and ask them whether they are more likely to use courts because of their physical proximity;

- Review court case files to see if low-level crimes have been adjudicated, as well as more serious offences (if people increasingly turn to the formal courts to resolve non-violent, low-level offences this may suggest an increasing confidence in the courts); and

- Interview local community leaders and criminal justice practitioners, such as judges, prosecutors and defence counsel, about the changes in access to justice that they would attribute to the project.

> **Proxy Measures**
>
> "Indicators are almost always proxies of the outcomes or concepts they measure. To varying degrees, indicators are removed and simplified from the outcome of interest in order to make it possible to measure them easily, frequently and at low cost. Their value lies in the fact that they are expected to correlate with the desired outcome, but the correlation is rarely perfect: changes in most indicators are fundamentally ambiguous."
>
> *Vera Global Guide, 2003, p.4*

This information about the perspective of different groups will help to understand both the effect that mobile courts are currently having on access to justice as well as ways to enhance the impact of these courts.

TABLE 2: CONCEPTS THAT ARE EASIER AND HARDER TO MEASURE	
Easier to Measure	**Harder to Measure**
Contact with the formal court system, such as the number of cases resolved by the courts in a given time period, using case file reviews or administrative data provided by the court.	Effectiveness of informal justice mechanisms — their operation may vary widely from place to place.
Gender balance in the police force (in countries where a reliable police census exists).	Actual sexual and gender-based violence rates over time; increases in reported SGBV cases may reflect an increase in the number of offenses, an increase in rates of reporting, or both.
Representation of racial, ethnic and religious minorities in the government.	Discrimination on the basis of sexual orientation. There is no reliable data available on this subject and survey questions may trigger confrontation and place people at risk of physical and emotional abuse.
Homicide rates, since homicides are usually reported UN or government partners are more likely to collect data on homicides compared to other offences.	Rates of non-violent crime, which may be rarely reported.
Perceptions of corruption through public or expert surveys. There are many corruption indicators available if this needs to be measured (see Appendix F).	Actual corruption, as only a small percentage of government officials are prosecuted for corruption and the public may be reluctant to acknowledge payment of a bribe out of fear of retaliation data on corruption is hard to come by.
Percentage of people in pre-trial detention. Most prisons and many UN agencies collect this information.	Reasons for lengthy pre-trial detention. This requires reliable data, which is often non-existent, from the police, investigating agencies, prosecution, courts and prisons.

SECTION 2: PLANNING MEASUREMENT

This section outlines key steps of the planning process. It describes ways to determine project goals, develop appropriate measures and decide whether to rely on UNDP internal research capacity or employ outside consultants. This section also provides basic guidance on hiring and working with outside consultants.

The planning process consists of four main steps and will inform decisions about how to best implement the measures, including whether a project has the necessary capacity in-house or will need to hire outside consultants.

Identify priorities and goals	→	Determine scope of measurement	→	Choose measurement approach	→	Determine skills required to conduct assessment
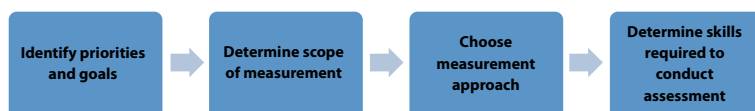
2.2.A. Identify Priorities and Goals

UNDP programmes and projects are always tied to UN Development Assistance Frameworks (UNDAF) and Country Programme Documents (CPD), which define the outcomes.

BOX 6: FORMULATING OUTCOMES AND OUTPUTS

Outcomes are actual or intended changes in development conditions that interventions are seeking to support.

Outcomes describe the intended changes in development conditions that result from the interventions of governments and other stakeholders, including international development agencies such as UNDP. They are medium-term development results created through delivery of outputs and the contributions of various partners and non-partners. Outcomes provide a clear vision of what has changed or will change globally or in a particular region, country or community within a period of time. They normally relate to changes in institutional performance or behaviour among individual groups. Outcomes cannot normally be achieved by only one agency and are not under the direct control of a project manager. Outcomes (or outcome level results) are normally defined in the UN Development Assistance Framework and the Country Programme Document.

Outputs are short-term development results produced by project and non-project activities.

Since outputs are the most immediate results of programme and project activities, they are usually within the greater control of the government, UNDP or the project manager. Outputs generated by projects are always connected directly to an outcome. There is a critical responsibility at each project level with regards to the generation of the planned output through a carefully planned set of relevant and effective activities and proper use of resources allocated for those activities.

FIGURE 2: OUTCOME-OUTPUT-PROJECT NEXUS
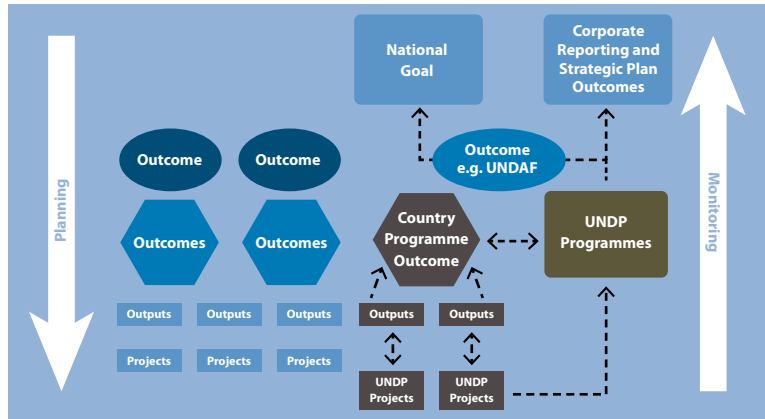
The very first step to establish a project's effectiveness is to define, and agree with stakeholders on, outputs and indicators for success. This must be done before the project is implemented. The process of collecting assessment data will help refine project outputs and establish the baseline data. For example, if the assessment phase reveals that one of the main reasons for excessive pre-trial detention is lack of capacity for investigation and case management, there is a need for relevant capacity building activities. The assessment can help to define clear baselines and collect necessary data. This will make it easier to identify the project outputs. Once project outputs have been identified, the design of output indicators as well as targets to assess progress as part of mid-term and final evaluations can commence.

The process of developing indicators and targets may require a rethinking of whether outputs are realistic and appropriate. For example, rather than focusing on the

TABLE 3: SMART OUTCOMES AND IMPACTS		
S	Specific	Impacts, outcomes and outputs must use change language — they must describe a specific future condition.
M	Measurable	Results, whether quantitative or qualitative, must have measurable indicators, making it possible to assess whether or not they have been achieved.
A	Achievable	Results must be within the capacity of partners to achieve.
R	Relevant	Results must make a contribution to selected priorities of the national development framework.
T	Time-bound	Results are never open-ended — there must always be a targeted end date.

Source: Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), Figure 10 on p.58.

broad goal of increasing access to justice, a focus on more concrete, measurable objectives such as "access to justice for victims of SGBV" or "enhancing confidence in local courts among the poor in a particular province of a given country" may be more productive. Although the ability to measure outputs should not influence programmatic areas, project outputs should be SMART, that is: specific, measurable, achievable, relevant and time-bound, to ensure that progress can be tracked (see Table 3, Box 13 and section 3.1.B.).

2.2.B. Identifying and Engaging Stakeholders

It is essential that RoL projects are owned by national stakeholders, including governments, civil society organizations and members of the public.[20] These stakeholders include individuals who "will benefit from the development activity or whose interest might be affected by this activity."[21] If they are fully supportive of the project's aims and methods, and made aware of plans to measure project impact, stakeholders can be important supporters and valuable allies. However, if they are unaware of measurement plans and have not been informed of the release of measurement findings then stakeholders may, understandably, feel marginalized. A (real or perceived) lack of stakeholder involvement will increase the likelihood that findings and recommendations will be blocked, derailed, or perceived as ineffective or threatening. Project staff may already know who the primary stakeholders are, especially if they have worked in the country for a long period of time. If this is not the case, stakeholders can be identified by mapping those whose work, interest, or problems will be affected by the project that is being measured.

Once stakeholders have been identified staff should develop a strategy to engage them as early as possible and throughout the course of the project to attain their involvement and support. To secure this support, conduct regular stakeholder briefings and provide project updates or other short documents describing the progress of the project along with any initial findings. A project advisory group is another way to engage stakeholders and incorporate their advice on project methods and how to translate findings into policy recommendations.

It is generally impossible, if not impractical, to engage every possible stakeholder, or everyone who may be affected by, or has a vested interest in, the RoL project. Prioritizing certain categories of stakeholders may be necessary for deciding who to engage when designing measures and sharing results.

---

20 See Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), Section 3.4 on p. 93 for additional guidance on engaging stakeholders in measurement. Also see Note on Assessing the Rule of Law, Justice and Security Sectors Using Institutional and Context Analysis and a Political Economy Perspective (UNDP, Draft 2012), Section 3, step 2 — Stakeholder and Engagement Analysis.

21 Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), p. 25.

Primary stakeholders include major decision-makers whose support is essential for the success of a project, and proponents or initiators of measurement activities. Primary stakeholders might include project partners, supervisors and funders. Individuals who are affected by a project such as inmates, arrestees, women, or militants are usually primary stakeholders. While it may be logistically impossible to engage all primary stakeholders it is important to identify and work with groups that represent their interests (e.g., community leaders or pastors who visit prisons). Secondary stakeholders are individuals whose support is important for data collection and the interpretation of findings. National governmental representatives, civil society partners, or other UN agency staff may be primary or secondary stakeholders, depending on their relationship to a project. Secondary stakeholders typically include those who are interested in findings because they operate in the same field. While these individuals can provide assistance (e.g., participate as experts in a survey or help to contextualize findings), they do not typically have enough influence over the structure and implementation of measurement activities to be considered primary stakeholders.

2.2.C. Determine Measurement Scope

Once outputs and activities of a project have been identified and agreed with stakeholders, it is important to define the scope of any measurement activity. For example, is a particular demographic group in focus? Will data be collected on everyone who enters police custody, or a particular subset of arrestees? Will measurement activities cover an entire country, or a selection of regions? These decisions will have a significant impact on the design of mid-term evaluations and final evaluations. For example, if the project's focus is on access to justice for women in a particular region, interviews with the male respondents may be unnecessary unless their experiences are being compared to those of women. Similarly, a project that addresses access to informal justice mechanisms in rural areas would not require interviews with city residents. The following checklist describes important considerations when determining the scope of measurement work.

Box 7: Weighing Risks and Benefits

In some cases, the risks associated with collecting information may outweigh the benefits. For example, conducting research interviews with members of sexual minorities may expose some of these individuals to a broader public and put them at risk of being physically and emotionally abused. In many instances, conducting anonymous research on sexual orientation and gender identity can minimize the risks and maximize the benefits of data collection.

1. Is the project targeting specific regions or sub-regions?[22]

2. Within project sites, will all people be affected equally, or does the project target particular social and/or professional groups?

3. How are the needs of women, children and other marginalized groups accounted for?

4. Who is the primary audience, and what types of research evidence will they find persuasive?

5. What information on resources and budgetary considerations is needed to determine the feasibility of expansion?

6. How will success be measured? Is it important to have quantitative measures of success?

22 If a project is set up with baselines limited to a certain region (e.g., Kivu in DRC, or Baluchistan in Pakistan), then subsequent evaluations will also focus on these areas.

2.2.D. Choose Approach to Measurement Steps

Once the scope of the project has been finalized, the next stage is to decide how to approach the three measurement steps. This could, for instance, involve deciding between conducting qualitative in-depth interviews, or a public survey (described in Chapter Three). Prioritizing information needs can help steer this decision-making process. For example, if a project aims to provide legal representation services to juvenile defendants facing criminal charges, it may be essential to know how many young people have been contacted by project lawyers, and the outcomes of their cases. Knowing the extent to which young people and their families are satisfied with the services that they receive may be an important, but lower priority, outcome to measure. A prioritized list of information needs can help sort out what scarce resources to allocate to data collection and analysis.

It is also important to identify the indicators that provide measures of progress that can be tracked over time. In order to identify these indicators start collecting data before the project is implemented to provide a starting point, or baseline measure. It will also be necessary to develop specific milestones and targets during the assessment. This will provide a way of gauging if the changes associated with the project are sufficient to claim project effectiveness. The following case study illustrates a specific indicator, baseline and targets.

Relevant Case Study	In 2010, UNDP provided targeted technical assistance to the government of Haiti to undertake measures to reduce pre-trial detention and strengthen the capacity of key criminal justice institutions to facilitate more efficient case management and better access to services for the most vulnerable populations.

Possible Measurement Approach

Indicator	Average number of days in detention (including police station, local jail and prison) for all pre-trial detainees in a specific prison.
Baseline	Average = 112 days (for 55 pre-trial detainees measured during assessment) • Median = 78 days • Minimum detention length = 2 days (3 detainees) • Maximum detention length = 475 days (1 detainee) • Percentage detained for more than 183 days (6 months) = 64% (35 out of 55 detainees) • Percentage detained for more than 365 days (12 months) = 13% (7 out of 55 detainees) • Comparison by gender, age, ethnicity, religion, and type of charge
Target Year 1	Average = less than 100 days (in month 12, measured at mid-term evaluation 1)
Target Year 2	Average = less than 80 days (in month 24, measured at mid-term evaluation 2)
Target Year 3	Average = less than 60 days (in month 36, measured at final evaluation)

The choice of measurement design will have significant resource implications, and decisions about what to measure should be scaled accordingly. When resources are scarce, a smaller number of stakeholders or project clients can be interviewed using qualitative research methods (see Chapter Three for more on qualitative versus quantitative measurement approaches).

Box 8: In-house vs. External

Hiring an external consultant may add greater objectivity and credibility to measurement activities. However, working with external consultants can increase costs and require a significant investment of time for the hiring process. In-house evaluations tend to be timelier, often less expensive, result in better buy-in and incorporate greater insider perspective. Even with an external evaluator it is important to know the basics of measurement (see Chapter Three) to develop an appropriate scope of work and manage the expectations of a funder.

2.2.E. Assess Skills Required

The skills required to collect, analyse and interpret data can vary substantially depending on the measurement approach. Designing and implementing measurement plans often requires a particular skill set, such as developing a sampling design for a national survey, conducting statistical analysis, or qualitative data collection with marginalized groups. If these skills are not available in-house it may be necessary to hire a consultant with the appropriate expertise. In addition to methodological expertise, qualified researchers should be able to apply their knowledge to crisis-affected and fragile settings.

It is often hard to find individuals with both substantive and methodological expertise so, in these instances, it is often best to assemble a team of researchers and project staff with the range of required skills and experience. When building a multidisciplinary team it may be important to include specialists from the international community, but it is also strongly advisable to include national experts who can provide necessary insights about the local context. Working with local consultants and organizations has myriad benefits. Partnering with local experts can not only enhance the design of a RoL project, it can also build credibility amongst national governments and grass-roots organizations. These partnerships ensure that research funding stays within the country and can increase the capacity of consultants and local research institutions to conduct similar studies in the future. Unfortunately, local expertise is not always available, particularly in crisis-affected settings, so regional organizations and experts can help to provide local context.

TABLE 4: POINTS TO CONSIDER WHEN SELECTING CONSULTANTS		
Expert Type	**Advantages**	**Disadvantages**
In-house	- Knowledge of local context and UNDP system - Proven record - No need to advertise and interview	- Time and scheduling constraints - Lack of methodological specialization - Organizational baggage
External	- Deep expertise in specific methods - Objectivity and positive perception by stakeholders - Fresh perspective	- Time and scheduling constraints - Lack of methodological specialization - Organizational baggage - May be unwilling to work on short notice - High compensation - Logistical challenges of travel and residence - May not know the local context

2.2.F. Decide Whether to Rely on In-House Expertise

There are a number of considerations when deciding whether to rely on in-house or external expertise (see Table 4).

When hiring an external consultant, project staff will need to follow these steps:

Decide to hire a consultant → Determine cost and availability of funds → Advertise position → Hire a consultant → Review proposed measurement design → Gauge challenges and provide support → Check-in on progress → Review deliverables
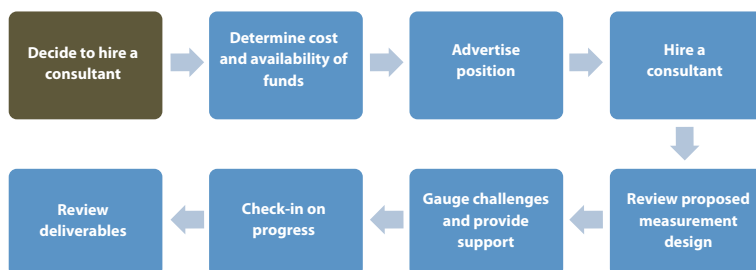
In the search for external consultants — from advertisements for the position to candidate interviews to development of candidates — it is best to be very specific about the required skills, expected deliverables and work timetables.

SECTION 3: MEASUREMENTS IN DATA POOR SETTINGS

This section offers practical guidance on how to navigate some of the time, budget, data, political and cultural challenges that arise when conducting measurement activities in situations that are data-poor. Measurement is often made more difficult by the degree to which a context is affected by crisis — indeed, crisis and fragility are often reasons why measurement is not carried out effectively. But while crisis heightens potential difficulties, challenges also arise in many disadvantaged and impoverished settings. While these challenges are intertwined, the following recommendations are divided into five thematic areas for practical purposes.

2.3.A. Time Constraints

This challenge can be the most pressing in situations affected by crisis and fragility. Throughout the process of measurement, and particularly during the assessment phase, two types of time constraints should be considered. The first relates to unavoidable delays, such as political unrest or severe weather, which may hinder data

TABLE 4: POINTS TO CONSIDER WHEN SELECTING CONSULTANTS		
Challenge	**Description**	**Possible Solutions**
Urgency of a project leaves no time for assessment.	Staff may have very limited time to plan a project given the immediacy of the assistance needed.	Start the assessment as soon as the project begins by collecting baseline data as soon as possible. This approach will allow staff to collect early warning data to help gauge whether project activities are appropriate and make the appropriate adjustments to maximize effectiveness.
Project activities are curtailed due to conflict or instability.	The timing of unforeseen events, ranging from post-election turmoil to a cholera outbreak, can significantly hinder data collection activities.	Restructure measurement activities to initially target areas least affected by the crisis, and survey those in most affected areas once the conflict subsides.
Evidence of success is required before data has been collected and analysed.	A new government or funding partner may ask for proof of a project's effectiveness before offering their support.	-Engage stakeholders in the measurement process from the beginning, and, if this is a new appointee, from the moment they assume office. -Make clear that premature measurements or predictors of success are unreliable and misleading. -Provide initial findings if they are available with a caveat that they are preliminary and, if necessary, confidential.

collection activities. The second constraint refers to the time required to complete data collection, analysis and interpretation. This distinction is important because while staff may have no control over delays, they can still design measurement activities in a way that ensures the timeliness of results. The timeline for a measurement exercise should be planned in conjunction with project needs and stakeholder interests.

2.3.B. Budget Constraints

As with any other activity, it is essential to carefully plan a budget before collecting data.[23] It is recommended that staff allocate 10 percent of the project budget for assessment, mid-term evaluation and final evaluation (possibly including the hire of a measurement expert for the project).

Despite the most careful planning, projects regularly exceed their budgets for a number of reasons. For example, certain services or resources may become more expensive, even within a short period of time. Project implementation delays can lead to measurement cost overruns. Escalating security concerns can increase the cost of data collection. When faced with budget constraints, limiting the scope of measurement activities may be unavoidable. Other options for overcoming budget constraints are outlined below in Table 6.

23   See Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), Table 19 on p.92 for key issues to consider when estimating the cost of an evaluation.

TABLE 6: STEPS TO OVERCOME BUDGET CONSTRAINTS		
Step	**Limitations**	**Possible Solutions**
Use secondary data, i.e., data collected by the government, civil society and international institutions.	It is often hard to assess the reliability of data when it is unclear as to how and why it was collected.	- Collect similar data from different agencies to compare data sets and sort through any discrepancies between data sources. - Disaggregate data into smaller units to assess their validity (e.g., compare spending for local courts in a specific region in a given month with paper records, or the views of experts).
Rely solely on local researchers for data collection, analyses and interpretation.	In some contexts, local researchers may lack necessary skills and training.	- Hire local research staff if they possess the necessary skills. - Provide training on data collection and management, computer spreadsheets and basic analysis techniques, if necessary. - Supervise data collection through systematic monitoring to identify and address emergent challenges.
Use volunteers wherever possible.	Volunteers may require training and, in some instances, may be less vested in the project than paid staff.	- Adopt strategies described above for local researchers. - Make sure volunteers know their work is valued.

TABLE 6: STEPS TO OVERCOME BUDGET CONSTRAINTS (CONTINUED)		
Step	**Limitations**	**Possible Solutions**
Supplement existing activities to collect data for new projects (e.g., add questions to surveys, or ask field staff to observe conditions in prisons or police stations).	Staff may be overstretched with current data collection and management tasks to take on new responsibilities.	- If collecting data for future projects proves to be too much just now, the data points can be part of planning for future projects. - Lessen the burden of data collection on staff by removing and replacing any data points that may be obsolete.
Identify local government officials to help collect and analyse data.	Government staff may not have the skills or training required.	- Identify national statistical offices, justice observatories or other existing government research divisions for possible staff who can assist with data collection. - Using government officials may at first require an initial investment of resources, but this can build capacity in the longer term.
Adopt the least expensive research design and data collection strategy.	Basic designs may be insufficient to gauge project effectiveness.	If there is no interest in the causal link between project activities and outcomes, do not collect data for comparison groups (i.e., similar groups that do not receive the service or intervention that is being tested).
Simplify existing designs as needed.	Making significant changes to a design may require updating stakeholders and retraining staff.	- Brainstorm possible revisions to a design with stakeholders and elicit their feedback on potential challenges. - Instead of in-person meetings, circulate a document to stakeholders describing changes and provide the option of follow-up meetings or phone calls, if needed.
Decrease the sample size or the geographic areas of data collection.	This may pose challenges in terms of obtaining findings that can be generalized.	- Make sure that samples are sufficiently large to produce meaningful findings. Small samples may not be reliable, especially when measuring change over time and comparing findings among different groups. - When selecting geographic areas, be sure to include regions that represent a range of demographics (e.g., urban and rural regions) in the sample.

2.3.C. Data Constraints

One of the greatest challenges of conducting measurement in crisis-affected and fragile settings as well as highly disadvantaged least developed countries, relates to the quality and availability of data. Data, as well as the resources and capacity to collect data, may be entirely unavailable or the information may be collected but incomplete. In other instances, political and security factors can compromise the quality of the data that is made available, or information may be of poor quality due to improper data collection and management skills within government offices or international agencies. Data may also be limited because of missing or incomplete information. Table 7 offers suggestions for overcoming some of the most typical data constraints, although the use of statistical techniques to address problems related to missing data exceeds the scope of this Guide.[24]

---

24   For techniques of dealing with missing data, see Allison, P. D. (2001) Missing Data Thousand Oaks, CA: Sage Publications. Also see Little, R.J.A. & Rubin, D.B. (1987). Statistical analysis with missing data. New York: Wiley.

TABLE 7: STEPS TO OVERCOME DATA CONSTRAINTS		
Challenge	**Description**	**Possible Solutions**
No baseline data available.	Without information on circumstances that predate a project, it is often impossible to gauge how effective an intervention is.	- Collect baseline data before launching a project, as part of the assessment. - Survey relevant respondents to see if they noticed any positive, or negative, changes as a result of the project (see Chapter Three for more).
Problems determining whether a data set is com-promised.	Outside of an obvious error or omission, it may be difficult to notice inaccuracies in mea-surements.	- Review all items to see if they make sense and correspond with what is known about the issue (i.e., face validity). - Speak with a colleague from a government partner to deter-mine if the numbers appear correct. - Compare findings with an independent data source (e.g., compare government data describing the size of a police force with United Nations Police data).
Data is compromised for political reasons, or due to a lack of capacity.	It is possible for an agency to falsify or selectively record data to make them look more favourable. On the other hand, some people or agencies that provide information have lim-ited data collection capacity so errors may be unintentional (e.g., because record books have been lost, or data was entered incorrectly).	- Use data collected from different agencies as a comparison (see above). - Verify data for a subgroup of cases (e.g., contact a sample of police stations to request information on the number of SGBV complaints received).
Security situation may hinder data collection.	Because of a war or civil unrest, sending researchers to the field may be dangerous. Access to conflict areas may be limited due to the destruc-tion of infrastructure or the unavailability of transporta-tion.	- Collect data for conflict zones by interviewing those who fled the areas. - Survey key informants remotely (e.g., over the phone, or once a person leaves the conflict area). - Request security support to ensure the safety of any research-ers in the field.
Data does not include un-derserved, or hard-to-reach populations.	Hard-to-reach groups may include SGBV survivors who have not reported their victim-ization, persons with HIV/AIDS, prison inmates, the disabled, religious and sexual minori-ties, or members of paramili-tary groups, rebels, insurgents and militias, particularly if they are children.	- Survey people, such as families, or community members or local researchers, who may know about the experiences of such populations. This methodology is known as indirect interviews. - Ask key informants from government, civil society and interna-tional agencies about the experiences of such populations. - Use snowball sampling techniques (as described in Chapter Three).

2.3.D. Political Constraints

Political constraints can greatly hamper measurement activities. For example, gov-ernment partners may be reluctant to collect and provide data, they may refuse to implement recommendations or acknowledge findings, and, in some cases, they may ask for an evaluation that portrays them in a favourable light and supports their political aspirations. Furthermore, while measurement activities are typically designed to inform a particular UNDP project or programme they may also reveal weaknesses or flaws in the work of national authorities; additional barriers to the adoption of recommendations can be introduced if the reputation of project stake-holders is at stake.

TABLE 8: STEPS TO OVERCOME POLITICAL CONSTRAINTS		
Challenge	**Description**	**Possible Solutions**
Government turnover and changes in the administration may undermine a project.	Given that government officials can change often, support from a previous administration may no longer guarantee access to data.	- Notify a new official about the project as soon as possible and explain its possible benefits. Do not wait until there is a need for data before making such a contact. If possible, arrange for a transition meeting that requires the outgoing partner to introduce the project to their successor. - Consider prioritizing data collection so that all the required data can be collected before the partner leaves office. - Sign a memorandum of understanding to institutionalize partnerships.
Obtaining government buy-in may not be feasible.	It may be difficult to secure and maintain cooperation because a partner is sceptical that measurement activities will improve existing practices.	- Understand the partner's priorities, data capacities and experiences in similar projects to ensure that requests are not viewed as unreasonable or overly taxing. - Establish common ground by emphasizing how measurement can help their agency. - If feasible, offer to collect additional data or produce reports that will be helpful to a partner. - Seek the support of senior UN officials for political dialogue and advocacy.
Releasing findings at election time.	The completion of a measurement study may coincide with an election that may jeopardize any dialogue about remedying problems.	- Release findings to involved parties on a confidential basis. - Produce a report for public dissemination after the election, and only after a partner has had the opportunity to comment.
Findings can be used for political manoeuvring.	Officials may inappropriately use findings to undermine the work of their political opponents or governmental or international actors.	- Keep all key stakeholders informed about the progress and findings of any measurement study so that they are updated directly and not from a politically motivated third party. - Be aware of possible third party obstruction when choosing a time, place and manner to disseminate findings.
Gauging informal justice mechanisms can be sensitive.	Measuring projects designed to alleviate human rights problems, or other challenges, related to informal justice mechanisms may alienate informal justice leaders (chiefs, elders and spiritual leaders) or their supporters who may refuse to provide data.	- Engage informal justice leaders as stakeholders in measurement from the beginning of a project and seek their advice on how to overcome these challenges. They may serve as gatekeepers to these systems. - Act as an observer and not as an advocate. Refrain from expressing a position on informal justice.

2.3.E. Cultural Constraints

The importance of understanding culture when conducting measurement activities cannot be overstated. Development agencies often rely on outside consultants to collect, analyse and interpret data. While these consultants may possess technical knowledge, their lack of awareness of the local context may limit their effectiveness. Speaking local languages, dressing properly, using appropriate gestures, and acknowledging respondents' efforts to provide data or help contextualize findings are all necessary steps for a successful measurement initiative. The following table provides suggestions for addressing some of the most common challenges relating to cultural constraints.

TABLE 9: STEPS TO OVERCOME CULTURAL CONSTRAINTS

Challenge	Description	Possible Solutions
Insufficient understanding of local context.	Surveys and other data collection techniques may be viewed as inappropriate, or use language that is either difficult to understand or not relevant to the local context.	- Rely on local expertise, or work with national experts who come from a project's target regions. - Pre-test data collection tools. - Make changes to data collection protocols if sudden problems arise.
Contacting female respondents may be difficult.	Cultural norms about the role of women may restrict the ability to recruit female respondents for focus group surveys and may impact the truthfulness of responses.	- Recruit female data collection staff to increase response rates among women. - Select interview places (e.g., private areas or places far from their communities) where women are more likely to feel comfortable participating in a survey.
Challenge of being interviewed by a stranger.	If discussing personal concerns is not considered culturally acceptable, some questions may cause discomfort and, in extreme cases, lead to confrontation.	- Ensure that a questionnaire includes a warning or prefatory question to evaluate respondents' comfort levels regarding sensitive questions. - Ensure that interviewers inform respondents of their right to refuse to answer questions and terminate interviews if they feel uncomfortable (see Appendix F for a sample Informed Consent form). - Conduct cultural sensitivity training for all research staff.
The value of responses may be impacted by respondents' perceived social status.	When reviewing findings, it may prove difficult to convince stakeholders that responses by the poor and vulnerable are as valid as those who hold high positions or belong to a specific socio-economic group.	- Explain the value of public survey data and how regular citizens possess a unique and valuable perspective on justice services or security based on their experience. - Break down findings by socio-economic characteristics to show possible differences in experiences or perceptions.

KEY TAKEAWAYS

Chapter Two—
What to
Measure?

• Develop appropriate measures for an RoL project
• Rely on proxies or multiple measures
• Plan measurement steps
• Identify project outcomes and outputs
• Assess the skillset required to conduct measurement activities and decide whether to hire an external consultant
• Respond to common data collection challenges in conflict-affected and fragile settings

HOW TO MEASURE?

CHAPTER 3

HOW TO MEASURE?

The methods described in this Guide will enable RoL project staff to use data to inform project design, overcome obstacles to project implementation and measure the effectiveness of a RoL intervention. Choosing how to collect and analyse data in order to measure impact, however, is a balancing act. On one hand, project staff will need to ensure measurement efforts are of sufficient quality to inform design and assess effectiveness. On the other hand, data collection must be feasible, especially in situations of limited resources and difficult conditions on the ground. In such circumstances it is worth remembering that even a modestly-sized data set, if carefully designed, collected and analysed, can provide a solid foundation for the development of RoL programming.

This chapter will lay out the necessary tools for measuring the effectiveness of a project. It provides an overview of common measurement approaches and the key steps for research designs so that staff can make informed decisions about how to measure RoL projects and programmes. It also includes guidance on additional references and resources on RoL measurement.

SECTION 1: MEASUREMENT APPROACHES

There are two overarching categories of measurement data: quantitative and qualitative. The former category refers to numerical descriptions such as percentages and averages, and the latter to information presented in narrative form (e.g., summaries of observations, first-hand accounts and descriptions of a process). Neither is harder or easier, or more or less valid than the other. Many measurement initiatives apply a mix of quantitative and qualitative methods, capitalizing on the relative strengths of each approach.

Qualitative measurement methods (QualMM) are often used to provide a nuanced description of issues that are complex, or not easily quantifiable. They are typically used to study a limited number of cases in detail. For example, when measuring projects that support transitional justice processes, such as prosecution, truth-seeking and reparation, it may be best to observe and describe truth commission meetings and criminal trials, and review biographies, newspaper articles, trial, medical and burial records and other documents. It may also be important to interview people involved in these processes as survivors, human rights violators, facilitators and observers.

> **Box 9: When to Use Qualitative or Quantitative Measures**
>
> Qualitative measurement methods can be used at every step of a measurement process but are best suited for the initial assessment phase when the scope of the problem and a project design are being decided.
>
> Quantitative Measurement Methods are more suited to the collection of baseline or benchmark data.
>
> Combining these two approaches typically produces more accurate and complete findings.

BOX 10: STRENGTHS AND LIMITATIONS OF QUALITATIVE METHODS

Strengths ⟵ Qualitative Methods ⟶ Limitations	
Useful for studying a small number of cases in depth (e.g., a single incident when several inmates escaped from one prison).	**Difficult to generalize to a larger number of cases and places** (e.g., prison conditions across the country cannot be extrapolated from conditions of detention in one prison).
Generates a contextual description (e.g., A2J problems in a specific rural area can be explained in terms of conflict, poverty, traditions, illiteracy and other factors relevant to local communities).	**Not typically useful for predicting future events** (e.g., an analysis of A2J for a specific time period cannot predict future challenges).
Findings described in narrative form can be understood by a wide range of audiences (e.g., interpreting a narrative description of SGBV trials does not require technical skills or specialized training).	**Lack of credibility with stakeholders** (e.g., without providing statistical findings on how many SGBV cases have been adjudicated and their outcomes, stakeholders may not find a description of a few cases compelling).
Provide an opportunity for respondents to define their experiences, beliefs, attitudes and needs in their own terms (e.g., survey respondents may want to explain what fairness in courts means to them instead of responding to predefined definitions and measures)	**Not amenable to comparison between respondents or settings** (e.g., because survey respondents may define fairness in courts very differently, it will be hard to compare the experiences of different groups).

BOX 11: STRENGTHS AND LIMITATIONS OF QUANTITATIVE METHODS

Strengths Qualitative Methods Limitations	
Useful when measuring change over time with a large group of individuals or cases, and/or producing findings that can be generalized from a relatively small group to a larger population of individuals or cases (e.g., findings based on 40 randomly selected police stations from across the country can be used to gauge the issues facing all police stations).	**Difficult to generalize to a larger number of cases and places** (e.g., prison conditions across the country cannot be extrapolated from conditions of detention in one prison).
Can be used to determine if a project caused a specific outcome (e.g., administrative data can be used to assess whether providing the police with vehicles decreased average response times).	**Need for large sample sizes and high level statistical skills to collect and analyse data, particularly if the goal is to determine if a project caused a certain outcome** (e.g., 1000 or more public survey respondents may be needed to generate meaningful findings, especially in order to break numbers down by gender, ethnicity, region, etc.).
Can be faster than collecting qualitative data, especially when relying on existing data provided by other agencies (e.g., data on crime reports before and after the implementation of a project).	**Risk that data from other agencies is unreliable, especially if the quality of data cannot be double checked** (e.g., without understanding possible political biases, data collection capacity and the purpose of data collection, repurposing secondary crime data provided by the police may result in an inaccurate assessment of reported crime).
Can allow the prediction of future events (e.g., if SGBV rates are collected for several years, then the reported SGBV crime rate for the next year can be predicted and planning for projects put in place accordingly).	**Narrow focus on statistical information without capturing the full picture of people's needs, experiences and perceptions** (e.g., official SGBV rates will not reveal anything about main causes of SGBV, experiences of survivors, and reasons for underreporting).

Quantitative measurement methods (QuantMM) usually rely on numerical or statistical data, whether collected through administrative data systems, direct observations or quantitative surveys. For example, if targeting the provision of legal aid services in a remote area, the percentage of defendants that were presented by a lawyer before the project was in place (i.e., baseline data) could be compared with the percentage with representation once the project has been operating for a year (i.e., follow-up data). Alternatively, a project could monitor the homicide rate to see if it decreased after hiring and training a team of newly qualified police officers. Many

questions are quantitative in nature, including the numbers of crimes committed or reported, people accessing services, new facilities constructed or people trained by an RoL project.

3.1.A. Measuring Progress Using Indicators[25]

Indicators are measures of development[26] that are used in UNDP programming to track changes over time relative to the intervention planned.[27] Within the results-based-management framework, UNDP uses three types of indicators:[28]

- Impact indicators
- Outcome indicators
- Output indicators

Indicators rely on baseline data collection before a project is implemented; without a baseline measure to compare against, there is no way of knowing how much change occurred over the period of a project. Baseline data should be collected during an assessment while follow-up data can be gathered during mid-term and final evaluations. Findings are most reliable when there are multiple follow-up data collection points, which also help to verify that any changes observed during the measurement process are due to the RoL project and not related to external factors (see section 3.1.B. 'Isolating the Impact of a Project'). In other words, as a project gradually increases the dosage of treatment, it should see a gradual improvement in outcomes.

> **What is an Indicator?**
>
> "Indicators are signpost of change along the path to development. They describe the way to track intended results and are critical for monitoring and evaluation."
>
> *UNDP Handbook on planning, monitoring and evaluating for development results. P.61*
>
> "Indicators are quantitative or qualitative variables that allow stakeholders to verify changes produced by a development intervention relative to what was planned. Quantitative indicators are represented by a number, percentage or ratio. In contrast, qualitative indicators seek to measure quality and often are based on perception, opinion and levels of satisfaction. … There can be an overlap between quantitative and qualitative indicators. Some statistical data or information stated with number can provide qualitative meaning."
>
> *United Nations Development Group, Results-Based Management Handbook. 2011, p.19*

Consider the example of a project on strengthening the capacity of investigators and prosecutors.[29] To measure its effectiveness, the first step is determining how to measure 'capacity,'[30] the second is collecting data on capacity before implementing

---

25 For more on indicators and grouping indicators into baskets, see The United Nations Rule of Law Indicators (OHCHR/UNDP, 2011), pp.1-5

26 See also section 2.1.B, Using Proxies and Multiple Measures of RoL Concepts, p.16

27 See also Results-Based Management Handbook (UNDG, 2011), p.19

28 See Handbook on planning, monitoring and evaluating for development results (UNDP, 2009), p.65

29 UNDP plays an important role in developing the capacity of investigators and prosecutors in a number of countries to address serious crimes, including conflict-related crimes, in a manner consistent with international standards. For more, see UNDP Global Programme Annual Report 2010 (UNDP, 2010) and UNDP Global Programme Annual Report 2011 (UNDP, 2011)

30 Project staff may be able to measure this capacity in terms of the ability of the prosecutor's office to process cases without delays. This could be defined as a percentage of all cases processed and forwarded to court within the timeframe specified by a statute, or by the average number of hours between arrest and indictment, and/or between indictment and criminal trial.

the project, and last is continuing to collect the same data as certain milestones are reached and the project completed. If the data on capacity shows gradual improvements at multiple stages of project implementation (i.e., more treatment = more effect), this may suggest that project activities are related to the improvements.

Why not use a Single Indicator?

"The single governance indicator which captures the subtleties and intricacies of national situations, in a manner which enables global, non value-laden comparison does not exist. Using just one indicator could very easily produce perverse assessments of any country and will rarely reflect the full situation."

Governance Indicators: Users' Guide, (2nd ed) 2007, p.12

Indicators can be based on the full range of data sources described in this guide (e.g., surveys of experts and members of the public, document reviews, observations, administrative data). A number of initiatives combine multiple methods to measure the delivery of services or the performance of institutions. This multi-method approach is particularly valuable in conflict-affected and fragile settings where a lack of available data and barriers to collecting new information may compel researchers to develop more creative ways of gathering information and cross-checking data quality.

Given that most RoL issues — such as transparency, fairness, access and responsiveness — are multi-faceted, they require multiple measures (see section 2.1.C.). Therefore it is advisable to use groups of complementary indicators, often called 'baskets of indicators' (see Box 12 below).

A number of existing United Nations initiatives have developed indicators to assess RoL issues. For example, the United Nations Rule of Law Indicators include groups

---

BOX 12: INDICATOR BASKET EXAMPLE

Basket: 'Access to Justice'

Indicator 1: Ratio of urban to rural residents who report they have access to courts (Public Survey)
　　　　Rationale: If the ratio is close to 1, this indicates that people have equal access to courts whether they live in urban or rural areas.

Indicator 2: Percentage of cases processed by courts that involve minor offenses (Administrative Data)
　　　　Rationale: An increase in the proportion of low-level offenses being handled by the courts suggests that courts are being used to resolve a wide range of problems, an indicator of increasing accessibility.

Indicator 3: Proportion of experts who believe that indigent defendants are represented at any stage of criminal proceedings (Expert Survey)
　　　　Rationale: Access to free legal counsel is a cornerstone of access to justice.

Indicator 4: Ratio of male to female victims that report crime (Public Survey)
　　　　Rationale: Women typically experience greater problems accessing justice compared to men so this measure provides a proxy for equality of access.

(baskets) of indicators assessing performance, integrity, transparency, accountability and capacity of criminal justice institutions.[31] Furthermore, toolkits produced by UNODC[32] and UNICEF[33] include measures relevant to the assessment of criminal justice systems. Finally, civil society organizations or government statisticians at the country level may already employ useful indicators for RoL projects and building upon these can help to strengthen local capacity.

Appendix G includes a list of existing performance measures, tools and guides. Using existing indicators that have been tested in the field can save a significant amount of time. However, there are several factors to consider before adopting existing indicators (see Box 13 below).

---

31 Note, however, that the UN Rule of Law Indicators are used to assess the overall performance of criminal justice institutions (including law enforcement, courts, prosecution, criminal defence and prisons) and are not designed to measure the effectiveness of individual projects.

32 The tools have been grouped within criminal justice system sectors, including: Policing; Access to Justice; Custodial and Non-Custodial Measures; and Cross-Cutting Issues.

33 See, for example, Toolkits on Diversion and Alternatives to Detention (UNICEF) or Manual for the Measurement of Juvenile Justice Indicators (UNODC/UNICEF, 2006).

BOX 13: CHECKLIST FOR DEVELOPING SMART INDICATORS

Specific (S)
- Are each of the output indicators able to capture the types of changes that are likely to occur in the period of the project?
- Are the indicators specific to the changes that the interventions expected to be produced?
- Do the indicators specifically capture the experience of vulnerable groups, such as rural populations living in poverty?
- Can the indicators that capture general experiences be disaggregated to isolate the experience of particular groups?

Measurable (M)
- Is it possible to collect the necessary data on a regular, continuing basis, particularly through simple and cost-effective means?
- Will the data collected specifically for output indicators be reliably accurate?
- Is the data available at reasonable cost and effort?

Attainable (A)
- Are the results in which the indicators seeks to chart progress realistic?
- Do those whose performance will be judged by the indicators have confidence in them?

Relevant (R)
- Is the indicator relevant to the intended outputs and outcomes?
- Are the indicators measuring outputs, not simply activities?

Time-bound/Trackable (T)
- Is the data available at reasonable cost and effort?
- Are the indicators likely to record progress toward the outputs?

Source: UNDP Handbook on planning, monitoring and evaluating for development results. P.6

3.1.B. Isolating the Impact of a Project

Although indicators may help detect change, they do not typically provide the type of information required to attribute those changes to a particular project, or group of projects. In many places, the proliferation of RoL initiatives complicates the process of determining whether a single programme or project is meeting its objectives. Because it is usually impossible to account for all factors influencing the outcome that is being measured, results should be described in terms of 'association', i.e., how two or more developments are correlated, rather than 'causality', i.e., how project activities led to an outcome. The following example illustrates the problem of causality more clearly.

Project Objective: Decrease in sexual and gender-based violence in a small locality

▼

Activity 1: Police trained in SGBV investigation (plus other activities)

Indicator 1: % of police trained in SGBV investigation (plus other indicators)

▼

Output 1: SGBV investigation improved (plus other outputs)

Indicator 1: % increase in confidence in police investigations of SGBV (plus other indicators)

▼

Outcome: SGBV rate decreased

Indicator 1: % decrease in SGBV rates (plus other indicators). Means of verification = public survey

▼ ▼ ▼

Project Claim: Police training resulted in greater awareness that SGBV is no longer a 'family matter' and should lead to the arrest of the perpetrator, as with other crimes. In response to the project intervention, indicators show that the police became more vigilant in detecting SGBV.	**Alternative Explanation 1:** Tribal chiefs began referring SGBV cases to formal courts, increasing rates of prosecution and subsequently reducing overall rates of violence (as was the case in Somalia in 2010).	**Alternative Explanation 2:** A USAID-funded daily talk show on local radio aired at the same time as the RoL project and had similar goals. The program raised awareness of SGBV and encouraged victims to seek help from the women's protection units of their local police stations.

In some cases, where UNDP is the only agency active in an isolated region or when a project has very specific objectives, correlating project activities with a positive change may be possible. If, for example, the project under review is solely providing vehicles to police in a remote area of a country, improved police response times in those areas can most likely be linked to this intervention. In other cases, it may be more important to understand the aggregate impact of many projects sourced by international and bilateral agencies than to isolate their individual impact. To use the earlier example, a determination that the collective work of UNDP, USAID and other development agencies active in a country improved police response to SGBV may be sufficient. The additional advantage of a coordinated approach to evaluation is the ability to pool measurement resources across agencies, reducing duplication and providing information to improve coordination of RoL programming across agencies. Furthermore, local government partner agencies may become frustrated and experience evaluation fatigue if several development agencies are working on similar projects and requesting similar datasets without coordinating efforts.

> **Box 14: Isolating the Impact of a Project — Prerequisites for 'Causality'**
>
> There are three rules of causality (i.e., whether a project caused an observed result):
>
> 1. Temporal Priority — change happens after an activity, not before.
> 2. Concomitant Variation — activities are correlated with outcomes (e.g., the more textbooks provided, the lower the percentage of illiterate children).
> 3. Elimination of Other Viable Explanations — no other factors, whether efforts by other agencies or socio-political or economic developments, could have possibly caused the change (outcome).
>
> Causality always implies correlation, although the converse is not always true.

If it is important to isolate the specific impact of a project, measurement approaches that determine causality are crucial. These typically require some form of comparison group to establish a counter-factual (e.g., the proportion of women who would report SGBV if the UNDP project did not exist). There are two main groupings of designs that can be used to measure the effect of a project. These are:

- **Experimental designs,** where participants are randomly assigned to a group that is affected by a project, or a comparison group that does not receive services. In order to compare outcomes, it is important to track both groups using the same methods. Experimental designs are the only measurement tool that can establish a direct causal link, with a high level of certainty, between project activities and desired changes.[34]

- **Quasi-experimental designs,** which include a range of methods that approximate random allocation and are used in settings where it is either impractical or undesirable to randomly deny people services. For example, it would be unethical to deny defendants access to legal representation in capital cases.

> **Box 15: Challenges of Experimental Designs**
>
> Experimental designs are a valuable tool for demonstrating impact but they can be costly, logistically challenging and often face ethical problems. It is important for project staff to be aware of the challenges involved in implementing these methods, even if they are not responsible for designing an evaluation. This awareness will help all staff manage donor expectations and justify the choice of research design to stakeholders.

---

34   In the absence of random assignment to project and comparison groups, any observed differences between the treatment and comparison group may be the result of pre-existing differences between these groups. Social scientists have come up with various ways of addressing this problem, including matching (e.g., Propensity Score Matching techniques) and statistical methods for assessing programme impact (e.g. Interrupted Time Series Analysis). A detailed description of these methods exceeds the scope of this Guide.

Each of these methods can be costly and often require advanced statistical skills. However, the expenditure on a rigorous evaluation can represent great value for money. For example, if piloting an A2J project in one jurisdiction, then ensure this project's effectiveness before deciding whether to replicate it across the country. In this instance, the benefits of comparison group designs far outweigh any costs.

SECTION 2: DATA COLLECTION AND ANALYSIS[35]

This section focuses on how to collect and analyse data. The sub-sections are organized by each of the six main modes of data collection — administrative data, public surveys, expert surveys, focus groups, document reviews and observation. Most of the information obtained from these data sources can — and should, wherever possible — be broken down by subgroups to understand how experiences and perceptions differ by gender, age, race, ethnicity, religion, residence type (urban or rural) and income. This approach will help to identify disparities so that project activities can target the most vulnerable and poorly served sections of society.

---

35   Also see Handbook on Planning, Monitoring and Evaluating for Development Results (UNDP, 2009), section '7.5. Evaluation Methodology', p.172, for more on data collection methods.

TABLE 10: MEASUREMENT METHODS AND THEIR APPLICATION			
Name	**Description**	**Suited to**	**Sample Questions**
Administrative data (See section 3.2.A.)	Quantitative information that describes the operation of government and other agencies and a wide range of social phenomena (e.g., arrest rates, government spending and population demographics).	- Settings where administrative records are reliably compiled. - Assessing change over time using official records. - Measuring changes in budgets, staffing levels, provision of services, or other readily quantifiable indicators.	- What is the rate of increase of courts deciding on environmental protection cases? - What percentage of the displaced population lacks a form of legal identity? - What is the average period of pre-trial detention for children? - What percentage of the vulnerable or marginalized received legal aid or paralegal assistance?
Public survey (See section 3.2.B.)	Dialogue between a researcher and respondents to generate information about a range of social phenomena.	- Generating information on the views, beliefs or experiences of large numbers of people. - Gathering information on views, experiences or beliefs. - Addressing sensitive topics or complex issues in depth. - Comparing the views, beliefs or experiences of different groups or sub-populations.	- Do respondents perceive the court system as free from bribery? - What percentage of male and female respondents seeks resolution of claims or disputes through informal justice providers? - What percentage of users of a land registry are satisfied with the office? - What proportion of residents in rural and urban areas consider their land and/or property tenure secure? - What proportion of informal workers experienced police harassment or corruption in the course of doing business?
Expert survey (See section 3.2.C.)	Dialogue between a researcher and individuals who possess specialized knowledge about the issue of interest.	- Gathering knowledge about issues which require specialized knowledge. - When public surveys are too costly or dangerous to implement. - When there is limited time for data collection. - Tracking change over time.	- Have there been delays in receiving police salaries? - To what extent is alternative dispute resolution/mediation helpful in addressing court backlogs with just and timely outcomes? - Are referrals between legal and justice and other health, social, educational and administrative services effective? - Was the training of prosecutors effective and did it have a positive impact on criminal case processing? - Do early access schemes to legal aid work effectively to ensure legal assistance at the police/investigation stage?
Focus groups (See section 3.2.D.)	Group discussions between researchers and stakeholders (usually experts on specific RoL issues or members of the public).	- Including a range of viewpoints relating to the same issue. - Assessing areas of consensus and divergence of opinion. - Generating suggestions for addressing challenges.	- How could the public image of the police be improved? - What is the best way to address the needs of internally displaced people in the aftermath of an earthquake? - What are the main capacity challenges faced by a local prosecutor's office? - Are decisions and judgments on family matters, including divorce, custody and inheritance, fair for both women and men?

TABLE 10: MEASUREMENT METHODS AND THEIR APPLICATION (CONTINUED)			
Name	**Description**	**Suited to**	**Sample Questions**
Document review (See section 3.2.E.)	Systematic examination of reports, official records and other documents (court records, crime registries, budgets, statutes and regulations, media reports, photographs, etc.).	- Assessing official policies or legislative protections. - Collecting information in settings where fieldwork is too costly or dangerous. - Assessing the adequacy of record-keeping systems. - Conducting historical research.	- Is there a policy covering conditions of detention for children? - Do crime registries include dates of arrest and charge? - Do court documents include a clear description of the legal matter and disaggregated data relating to the parties? - Do budget documents identify expenditure by agency and month? - How are government or justice agents portrayed in the media?
Observation (See section 3.2.F.)	Attentive watching and note taking (obtrusively or unobtrusively) on services provided, the nature of interactions, and features of the environment.	- Gathering information on non-sensitive activities that can be observed in public. - Assessing compliance with standards. - Measuring the impact of capital projects.	- Has a project led to improvements in the availability of police vehicles? - Are civil registry offices/administrative licensing offices fully staffed and providing non-discriminatory services? - Are female detainees fully separated from men? - Are women treated differently from men during informal justice hearings? - Do police officers wear badges or other forms of visible identification?

3.2.A. Administrative Data

Administrative data (AD) includes a range of information collected by agencies or individuals, typically for purposes other than conducting research.

AD EXAMPLES
- Number of SGBV cases reported to the police in one month.
- Number of people who received paralegal assistance in one month.
- Number of cases processed by a local first instance court in twelve months.
- Number of inmates in pre-trial detention in a specific prison.
- Amount of funds provided for court administration for a fiscal year.
- Average delays in paying salaries to prison staff.

Collecting data

Administrative data can be self-selected by reviewing files, report books or other types of written records — an approach known as a case file review. For example, data on the number of firearm-related deaths can be gleaned from local hospital records, or the incidence of gender-based violence can be found in the occurrence books of local police stations. The benefit of a case file review is that one can format the data to suit the needs of a project and collect the most recent data available, which may not be accessible to other sources as yet. Case file reviews, however, can be costly and the scope of data collection may be limited.

BOX 16: JUDICIAL MONITORING IN EASTERN DEMOCRATIC REPUBLIC OF CONGO (DRC)

Process: When UNDP started programming on access to justice for victims of SGBV in Eastern DRC, little administrative data was available on the overall response by the judicial system in the Eastern provinces to the sexual violence cases. Consequently, in the framework of the UNDP project on Access to Justice and under the auspices of the Ministry of Justice and Human Rights, UNDP recruited a team of 10 national officers and trained them as 'judicial monitors' to collect data in all jurisdictions on each phase of the judicial process with regard to cases of SGBV in North and South Kivu provinces as well as in Ituri district in Eastern DRC. For the first time ever, the annual reports produced by the judicial monitors provided precise and reliable quantitative data on the treatment of SGBV cases at all levels of the penal chain. In addition, based on a sample number of cases, a qualitative analysis with regard to the respect of the rights of the parties to the process was provided. The main objective was to provide stakeholders with reliable data on the judicial response to SGBV and to raise the awareness of national authorities to the importance of making such data available to interventions designed to strengthen the capacity of penal chain institutions.

Results: The data and analysis of the judicial monitoring informed and guided UNDP's and other partners' programming on access to justice and in particular on access to justice for victims of SGBV. Furthermore, the judicial monitoring has proven to be an important tool for state-building efforts in DRC as it provides relevant information on the capacity and performance of the system and of the framework for quality control of its actors. Judicial monitoring helps identify resource needs and guide efforts supporting legislative and institutional change.

Using secondary data

Many governmental agencies and non-governmental organizations routinely collect administrative data, even in places with limited resources. For example, ministries of justice, civil society organizations or other development partner agencies may centrally record the number of people seeking legal and paralegal assistance, the number of inmates in local prisons, the number of police officers trained to respond to SGBV, or the salaries of judges and magistrates. Additionally, national bureaus of statistics (or their equivalent) may collect demographic information that is useful for measuring the effectiveness of a RoL project.

If an agency is not required to collect administrative data as part of its daily operations, it may be possible to trigger its data collection activities — and help improve its capacity to collect and use data — by providing training and technical support. Such partnerships can help improve the quality of administrative data while also enhancing local ownership over a project and its measurement. Note, however, that data that is essential for the operation of an agency is invariably more accurate than data that is collected solely for research purposes. Dedicated data collection activities will typically fail in the mid- to long-term in the absence of additional resources to ensure sustainability.

In addition to government agencies, a wide range of national, bilateral and multi-lateral actors collect administrative data. For example, the International Committee for the Red Cross and Médecins Sans Frontières record information on mortality and morbidity, often including data collected from correctional facilities. Similarly, other UN agencies, such as the Department of Peacekeeping Operations (DPKO), collect a wide range of administrative data on criminal justice systems. Mapping these data collection activities and building measurement strategies upon existing data collection systems will help preserve resources and may galvanize support among stakeholders.

How to Collect Administrative Data	
Primary Administrative Data Collect original administrative data for	**Secondary Administrative Data** Use data already collected by another agency for a different project/purpose.
BASIC STEPS 1. Determine what type of data is needed for what geographical area and time period. 2. Identify individuals and agencies, including NGOs, who can assist in data collection. 3. Obtain necessary permission for data collection. 4. Develop data collection worksheets to ensure consistency across data collection sites. 5. Hire and train data collection staff. 6. Create a protocol to ensure confidentiality of individuals referenced in data (as relevant). 7. Validate data by comparing it against relevant data collected by other agencies (where available).	**BASIC STEPS** 1. Determine what type of data is needed for what geographical area and time period. 2. Identify individuals and agencies who possess data. 3. Explain project aims. 4. Obtain necessary permission to receive data. 5. Request data (broken down by weeks/months and towns/regions, as this makes it easier to validate). 6. Check for computational and other mistakes and make sure data refers to the time period and area requested. 7. Request similar data from a second (or third) source as a check on data quality.

Additional resources on administrative data:

• This Guide, Section 2.3.C. for other tips on how to improve data quality

• DPKO/OHCHR RoL Indicators — '2.2. Accessing Existing Data', p.25, and '2.5. Accessing the Data', p.29

• UNDP Handbook on Planning, Monitoring and Evaluating — '7.5. Evaluation Methodology', p.172

• Measuring Progress toward Safety and Justice — '2.1. Using Administrative Data', p.7

3.2.B. Public Surveys

Public surveys can provide valuable insights on a range of RoL issues that may be impossible to measure using other data collection techniques. These can include so-called 'household surveys', which are a major source of social and demographic statistics in some countries, alongside population and housing censuses and the administrative record systems. Household surveys can be used for the collection of detailed and varied socio-demographic data, and in countries where these surveys are carried out it is worth exploring whether questions relating to the rule of law, justice and security could be added.[36]

Given the goal of UNDP projects to support quality of life improvements for all people, and particularly for the marginalized and vulnerable, the importance of assessing people's experiences and perceptions cannot be overstated. Surveys can be administered in-person, or remotely by phone, mail, email, or by using web-based software. However, in-person surveys are the best option in places with limited access to technology.

Surveys can adopt a range of qualitative and unstructured, or quantitative and structured questions (see Box 18). Structured questions include a list of possible

---

36 See for instance: http://unstats.un.org/unsd/demographic/sources/surveys/Handbook23June05.pdf

BOX 17: SAMPLING DESIGNS

Sampling is a strategy for selecting survey respondents in a way that guards against bias and helps ensure that the selected group is representative of the wider target population of interest (e.g., the general population, or all people living in a particular city or village).

Probability: all participants have an equal chance of being selected in a sample

Random sample	The basic design of a random sample is that each case is chosen randomly and entirely by chance, and everyone in the wider population has the same probability of being selected.
Systematic	Using a random start (i.e., the first case is randomly selected) and then every nth case (e.g., 5th, 20th, 100th) from a list of eligible cases is included in the sample (particularly useful for selecting a sample of people from official criminal records, e.g., every 10th arrest record, or every 5th prison record).
Stratified	Used to ensure representation of important groups. First cases are divided into groups (or strata) based on a variable of interest (e.g., race, gender, income) and cases are randomly selected from within each strata. Cases may be selected from each strata proportionately (same ratio as in the wider population) or disproportionately (e.g., oversampling minority groups).
Cluster	Used to address resource issues when conducting surveys over large geographic areas. After dividing eligible cases into groups (e.g., cities or states), a specified number of these groups are randomly selected. Cases can then be randomly selected from the populations in these clusters.

Non-Probability: some participants have no chance of being selected in a sample

Convenience (Accidental)	Cases are selected because a researcher has easy access to them and is not concerned with develop-ing a representative sample (e.g., interviews of people on the street).
Purposive (Judgment)	Cases are selected based on researchers' judgment and needs (e.g., survey participants are chosen because of their specialized knowledge; yet the sample does not necessarily represent all individuals with such knowledge).
Snowball	Each respondent is selected based on the referral provided by a previous respondent; researchers start with a few respondents and gradually increase a sample based on referrals (useful for reaching hard-to-identify and unknown groups, such as religious or sexual minorities).

Box 18: Examples of Qualitative Versus Quantitative Questions

Qualitative (unstructured)	Quantitative (structured)
How would you describe your interaction with the police? [Narrative response]	Were you treated with respect by the police? -Yes -No -Don't Know
Can you explain what influenced your decision to report a crime? [Narrative response]	Which of the following was the primary reason for reporting a crime: -To seek justice -To punish perpetrator -To protect yourself -Other, please explain…
Please explain the procedures for filing a complaint and pro-viding testimony. [Narrative response]	Were you able to file a complaint in a private area that was beyond the earshot of others? -Yes -No -Don't Know

responses from which the interviewer asks the participant to select the answer that most closely reflects their experience. When asking qualitative questions, the interviewer records the response in the participant's own words. Qualitative questions are generally suited to exploratory research, where it is not possible to predict the full range of responses when designing a questionnaire, or when addressing sensitive or complex topics (see Appendix E for a more detailed discussion of the relative merits of structured and unstructured approaches).

<div style="background:#005288;color:white;padding:1em;">

Box 19: Participatory Action Research (PAR)

Given that the success of measurement largely depends on building local ownership of activities, services and findings, project staff may choose to implement PAR to empower individuals who are affected by the project, build shared ownership and draw upon local expertise to inform research.

For more on PAR see: *Developing and Sustaining Community-Based Participatory Research Partnerships: A Skill-Building Curriculum*

</div>

Public surveys can be costly, especially if information needs to be collected for a nationally representative sample or when collecting data over large regions with geographically dispersed populations. Before embarking on an expensive and time-consuming survey, it is best to determine if there are similar surveys that already exist and are routinely conducted or will be conducted by national actors, UNDP or other agencies. These existing surveys may provide the baseline for an assessment. It may also be possible to add questions to existing public surveys. For example, if national actors or other UN agencies conduct routine local or national surveys, they may be willing to include additional questions. It is usually preferable to work with local polling organizations when conducting surveys.

The following steps are crucial when commissioning or conducting a survey from scratch:

1.	Identify scope of survey	Decide the issues being addressed and the intended beneficiaries.
2.	Develop questionnaire	Break down broader issues into more specific questions to ensure respondents understand questions consistently; translate the questionnaire into required languages and back-translate them to ensure accuracy.
3.	Pre-test questionnaire	Pilot survey questions with a small group of individuals from target populations.
4.	Develop sampling design	Determine a method for selecting respondents (see Box 17 on sampling designs).
5.	Hire and train data collection staff	Identify local NGOs and academics/students who can help to implement a survey; train interviewers to ensure consistency of data collection; create teams of data collectors and identify a team leader responsible for data collection and the delivery of completed questionnaires.
6.	Develop and implement data collection plan	Determine how data will be collected and recorded. Have data collection protocols in place to ensure the safety of research staff and the confidentiality of respondents, among other considerations (e.g., use informed consent forms for participants, see Appendix F).
7.	Analyse data	Consider quantitative, qualitative or mixed data analytical techniques as relevant (see section 3.1.).

BOX 20: PUBLIC PERCEPTIONS OF PALESTINIAN JUSTICE AND SECURITY INSTITUTIONS

Process: In 2011, UNDP set out to measure the impact of its A2J and RoL projects through a public perception survey which explored the following issues:
- Contact with justice and security institutions;
- Challenges to accessing justice and security institutions;
- Satisfaction with justice and security institutions; and
- Confidence with justice and security institutions.

The questionnaire was designed in consultation with national and international partners, and University College London, after which it was pilot tested and refined. In the summer of 2011, the survey was conducted by two national implementing partners working in close coordination with UNDP. The survey included 6,710 households.

Results: The findings of the survey suggested that investments in justice and security institutions were paying dividends. The data collected clearly showed that the overwhelming majority of Palestinian households believe that rule of law institutions were legitimate and chose to use them to resolve all manner of disputes: 91.7% chose to call the police when in danger; 71% considered that courts were the only legitimate institutions through which to resolve disputes; 63.3% were confident that they would receive prompt police assistance; 51.2% were confident that they could solve a civil dispute fairly through the courts; and 47.7% are satisfied that the public prosecution maintains dignity and human freedom. Such responses indicated that the recently created Palestinian National Authority justice and security institutions had already established their value. Survey data also revealed gaps in the occupied Palestinian territories' justice and security institutions. For example, factors contributing to the gap between women's and men's access to justice, and perceptions that the formal justice system is too slow. The data collected also yielded useful information for addressing such challenges. Finally, an extensive analysis of data enabled the team to derive recommendations for further strengthening UNDP's justice and security assistance in the occupied Palestinian territories.

BOX 21: PEOPLE'S PERSPECTIVE ON ACCESS TO JUSTICE SURVEY IN LAO PDR

Process: The Lao Bar Association, together with the Ministry of Justice and UNDP, commissioned a survey to gauge a representational cross-section of perspectives of Lao people in four provinces on justice and their interactions with the justice system. The 'Access to Justice' field survey occurred simultaneously in the four provinces and was conducted during a time that respected the agricultural calendar in order to avoid conflict with seasonal peak labour demand (such as clearing of the forest, or the planting or harvesting period). Each team was composed of one representative from the local Department of Justice, two ethnic researchers/facilitators, two students from the National University and two interpreters. This allowed for the survey to be undertaken in seven minority languages used in the four target areas. The researchers (one senior and one junior, one male and one female) had extensive experience in participatory methodologies, community development and facilitation. Their participation allowed for the cultural and linguistic bridging of the tools and concepts and ensured the accurate capture of local perceptions.

The tools and methodology used were developed in collaboration with civil society organisations, UN agencies and the Ministry of Justice. At the community level the tools used to carry out data collection included: interviews with the members of 24 village committees and 24 village mediation units; semi-structured interviews with 38 service users; 130 gender-segregated focus group discussions; and 600 individual interviews. The teams spent 4-5 days completing surveys in each village.

Results: The primary objective of the survey was to gauge a representational cross-section of perspectives of Lao people in four provinces on justice and their interactions with the justice system, and to collect empirical evidence about access to justice across the country. The survey was the first of this kind and provided policymakers, the legal system, civil society and development partners with a snapshot of the current capabilities of the Lao people to seek and obtain remedies for grievances, including family conflicts, violence, theft, land disputes, debt and other issues. The survey also informed and validated planned activities or implementation strategies, including those of the Legal Sector Master Plan.

Additional resources on public survey:

- This Guide, Section 2.3.C. for other tips on how to improve data quality

- DPKO/OHCHR RoL Indicators, '2.3. Collect Your Own Data', p.26, and '2.5. Accessing the Data, p.29

- UNDP Handbook on Planning, Monitoring and Evaluating, '7.5. Evaluation Methodology', p.172

- Measuring Progress toward Safety and Justice, '2.2. Using Survey Data, p.8

- Afro Barometer website, 'Survey and Methods' (sampling principles, interview methods, questionnaires, survey manuals), www.afrobarometer.org

3.2.C. Expert Survey

Expert surveys are an important source of data for a number of reasons. First, they generate in-depth information on technical or specialized issues that may be unfamiliar to members of the general public. For example, if a project is interested in payroll administration in relation to delays in salaries for the national police, staff will need to seek individuals who work for the payroll department or are recipients of salaries. Second, because expert surveys are not designed to be representative of a wider population, they often rely on a relatively small pool of respondents (20-30 people is not unusual) and are typically quicker to implement and less costly than large public surveys. Similar to public surveys, expert surveys can adopt a mix of structured or unstructured question formats, depending on the topic of interest.

WHO ARE EXPERTS?

Experts are individuals who have specialized knowledge. In the context of rule of law they may include high-level government officials, academics, legal experts, development agency staff, or regular citizens who have direct experience of an issue of interest (e.g., prison inmates, court users, hospital patients, militants).

There are a number of caveats governing the use of expert surveys that are important to consider. First, the criteria for defining 'experts' may be ambiguous and can depend on the perspective of different stakeholders. For some, experts are individuals who hold high positions in the government or have a detailed knowledge of an issue based on academic study. For others, experts can include anyone who has specialized knowledge of the issue at hand, irrespective of their professional position or affiliation. Unless a sufficient number of experts respond to survey questions, it may be impossible to report quantitative results. For example, if a project interviews 25 experts and only 15 of them provide valid responses (others either chose 'I don't know' or refused to answer), it will not be possible to draw general conclusions from the results. I may also be misleading to report results in percentages as this can mask the small sample size. There should be a minimum of 40-50 valid expert survey responses before referring to any related findings in terms of percentages.

Finally, expert surveys are particularly vulnerable to the perspective or affiliation of respondents. For example, if a project surveyed 20 experts from human rights NGOs about the prevalence of torture in national prisons, their responses may be entirely different from an identical survey of 20 government officials. To account for the variety of opinions that exists on most RoL topics it is important to include experts that represent the full range of opinions. Including multiple perspectives will also help to ensure that survey results are viewed as credible by a wide range of audiences.

Before collecting data for an expert survey, first determine a sample of expert respondents, identify areas of interest then create and pre-test a questionnaire.

Box 22: Steps for Conducting Expert Surveys	
Step 1	Select experts to participate and obtain their contact information.
Step 2	Identify languages spoken by the experts and translate questionnaires as needed.
Step 3	Identify the number of staff needed for interviewing experts.
Step 4	Train interviewers to ensure that they have a solid understanding of the project as well as good interviewing techniques.
Step 5	Address logistical needs such as transportation, accommodation and payment.
Step 6	Provide interviewers with questionnaires in appropriate languages and instructions on when and how to return both completed and uncompleted questionnaires.
Step 7	Collect questionnaires, assign codes to each expert, separate identification sheets from actual questionnaires and store the identification sheets in a locked cabinet. Staff will need to do the same with the questionnaires after data have been entered into computer files. Make sure that identification sheets and questionnaires are stored in two different locked cabinets.
Step 8	**Enter data into statistical or spreadsheet software.**

The identification of experts can vary but the most common sampling strategy is known as 'snowball sampling' (see Box 17). Using this method, an initial list of experts is asked to recommend others who possess relevant knowledge, who are then interviewed and asked to recommend further participants following an iterative process until the desired sample size is reached. To measure change over time, it is usually desirable to follow the same group of experts. This approach will help to ensure that surveys measure real changes and are not biased because of differences in the underlying beliefs or opinions of successive groups of expert respondents. Be aware, however, that interviewing the same individuals over time is not always possible because of, for example, high UN and government staff turnover and high job turnover among experts. This may be a particular problem when working in conflict-affected and fragile settings. If it is not possible to contact the same group of people when repeating surveys, try to maintain the balance of experts across various fields (e.g., maintaining the same relative proportion of academics, government personnel and NGO representatives).

Box 23: Confidentiality and Informed Consent

Whether conducting surveys or focus groups, project staff must ensure that data collection is voluntary, anonymous, or confidential, if necessary, and does not pose undue risks to respondents. Study participants need to be well informed about the possible risks and benefits of participating in a survey before they agree to answer questions (see Appendix F for a sample Informed Consent Form).

Additional resources on expert surveys:

- This Guide, Section 2.3.C. for other tips on how to improve data quality
- DPKO/OHCHR RoL Indicators, '2.3. Collect Your Own Data', p.26, and '2.5. Accessing the Data', p.29
- UNDP Handbook on Planning, Monitoring and Evaluating. '7.5. Evaluation Methodology', p.172
- Measuring Progress toward Safety and Justice, '2.2. Using Survey Data', p.8

Box 24: Characteristics of Focus Groups

Participants

- Selected based on expertise/experience
- Share common characteristics
- Ideally include five to eight participants

Moderator

- Ask questions to facilitate discussion
- Be an active listener and build rapport
- Display good management skills
- Remain neutral yet involved

Meeting Place

- Comfortable
- Circular seating
- Private area (e.g., room or office)

3.2.D. Focus Groups

Focus groups are one of the most commonly used qualitative data collection methods. They are usually arranged as a facilitated group discussion and typically adopt a semi-structured or unstructured questioning format, allowing members of the group to express their opinions in an unconstrained manner. Groups typically include five to eight participants, a moderator and an assistant moderator. The moderator is responsible for guiding the conversation and asking supplementary questions to follow up on topics of particular interest. The assistant moderator is responsible for keeping written notes (if the conversation is not being recorded) and asking additional questions as necessary. Participants should share a common background (e.g., all women, researchers, criminal justice professionals) so they feel comfortable expressing their opinions. If, for example, members of the public were asked to reflect on the performance of the police or judiciary in a mixed setting, they may not speak freely. Researchers either audio-record the conversation (with the permission of participants) or take meticulous notes.

Focus groups typically produce qualitative, narrative accounts that can be analysed by looking for common themes in participants' statements. They are particularly useful for assessing the diversity of experiences and are generally more cost-effective than one-on-one interviews as they allow multiple participants to be included in one session. Focus groups are usually not appropriate for sensitive or taboo issues, or on topics that could leave participants uncomfortable sharing experiences in a group setting. For example, one-on-one interviews or surveys would be preferable when assessing rates of domestic violence or other similarly sensitive topics. Because participants are asked to express their views and experiences in an open setting it is particularly important to provide clear information on topics to be discussed and to mitigate potential risks in advance.

Focus groups can serve as a potential data sources and a source of guidance for project design and review. They also provide a forum for pre-testing survey questionnaires, contextualizing and interpreting findings, and eliciting practical suggestions about how to use measurement findings to improve policy and practice.

Below are the key steps for establishing a successful focus group, however diverse.

1.	Determine the topic	Decide the purpose of the group and write 8-10 questions to guide the discussion.
2.	Identify and engage participants	Select up to eight participants.
3.	Recruit and prepare a moderator	Engage an impartial moderator who is skilled at facilitating meetings and knowledgeable about the topic. Choose an assistant moderator/rapporteur and train him/her.
4.	Select a meeting place	Arrange the meeting in a conveniently located private room where participants can sit in a circle.
5.	Hold a meeting	Lead the discussion by asking questions and making sure all participants understand them. Allow all participants to express their views.
6.	Record responses	Use audio/video recording or take meticulous notes (preferably by an assistant) while remaining engaged.
7.	Analyse data	Summarize findings by looking for common themes, preparing reports and presentations as needed.

Additional resources on focus groups:

- UNDP Handbook on Planning, Monitoring and Evaluating, '7.5. Evaluation Methodology', p.172

- Measuring Progress toward Safety and Justice, '2.2. Using Survey Data', p.8

- Guideline for Conducting a Focus Group, 2005, Elliot and Associates, http://assessment.aas.duke.edu/documents/How_to_Conduct_a_Focus_Group.pdf

- Designing and Conducting Focus Group Interviews, 2002, University of Minnesota, http://www.eiu.edu/~ihec/Krueger-FocusGroupInterviews.pdf

3.2.E. Document Reviews

Document reviews can include a wide range of materials, including court records, police crime registries, vetting documents, budgets, fiscal reports, written accounts of spending, newspaper articles, monographs and autobiographies, and pictures of accidents, people or corpses. Document reviews can help determine whether governments have provided sufficient information to the general public to ensure transparency and accountability (by publishing budgets, or information on the outcome of official investigations of corrupt practices, for example). Document reviews can also assess whether laws and regulations are consistent with international

Box 25: De Jure Measures

Given that the success of measurement largely depends on building local ownership of activities, services and findings, project staff may choose to implement PAR to empower individuals who are affected by the project, build shared ownership and draw upon local expertise to inform research.

For more on PAR see: *Developing and Sustaining Community-Based Participatory Research Partnerships: A Skill-Building Curriculum*

human rights principles or other practice standards. However, remember that having laws in place does not guarantee their implementation, and it is usually advisable to combine document review measures with other data sources, such as expert surveys or public surveys, to understand both the adoption and implementation of laws.

Below are the key steps for undertaking a document review:

1.	Determine the topic	Define the issue of interest (e.g., the content of police arrest records, or the extent to which corruption cases are reported in the national media).
2.	Identify documents	Determine the range of documents that include desired information and select a sample of documents (e.g., all arrest records, or only those in a sub-sample of police stations).
3.	Recruit and train reviewers	Decide who will conduct the reviews and provide training on specific elements of interest/standards for assessing documents.
4.	Create worksheets	Develop standardized templates for collecting information. If multiple researchers are conducting the review, then this will ensure consistency.
5.	Negotiate access	If documents are restricted and require access to specific settings (police stations, courts, archives), it will be necessary to provide all reviewers with documentation to allow them to access these settings.
6.	Collect data	Commence data collection and monitor the information collected; provide additional training as needed.
7.	Analyse data	Analyse data by looking for common themes and summarizing the content either quantitatively (e.g., percentages, averages), or qualitatively (narratives).

Additional resources on document review:

•   UNDP Handbook on Planning, Monitoring and Evaluating, '7.5. Evaluation Methodology', p.172 •   Measuring Progress toward Safety and Justice, '2.3. Using Narrative Reports', p.10

•   DPKO/OHCHR RoL Indicators, '2.2. Accessing Existing Data – Document Review', p.25

3.2.F. Observation

Observations can be conducted of criminal trials, prison conditions, police detention cells, informal justice proceedings and interactions between the police and members of the general public, among many other possibilities. Observations can be recorded through meticulous note-taking or by filling out observation worksheets once the observation has been completed if note-taking is too intrusive. In some settings, and with the necessary permissions, it may be helpful to take photographs. While most observations result in narrative summaries (QualMM) project

staff can also conduct an observation of a large number of institutions and produce quantitative results. An example of this would be visiting all prisons in a large country to assess the availability and quality of toilets and other basic sanitation systems.

The format of observations will depend on the nature of a project, but they will almost always require the following elements:

1.	Determine the topic	Decide what issues are being observed (e.g., how victims report crimes, how a tribal chief elicits testimony, how prison inmates cook and share food).
2.	Identify the location	Choose a place to conduct an observation (e.g., a police station, a court room, a prison, a village, a city square).
3.	Recruit and train observers	Engage an observer who is skilled in conducting similar observations, has knowledge of the topic and is likely to be non-intrusive and objective.
4.	Create worksheets	If staff are collecting similar data in multiple places and know what information to record, develop data collection worksheets to ensure consistency (this is particularly useful for quantitative data).
5.	Take notes	Take notes but allow sufficient time to observe and listen. It is possible to take notes after the observation if there is insufficient time or no opportunity to unobtrusively take notes during the observation.
6.	Repeat the observation	Repeat this activity in every place or individual location under observation (e.g., in every customs office, court house, prison, village, public transportation setting of interest).
7.	Analyse the data	The analysis method will depend on the type of observation (e.g., project staff may want to report percentages if there are enough cases, narratives, or both).

Additional resources on observation:

- UNDP Handbook on Planning, Monitoring and Evaluating, '7.5. Evaluation Methodology', p.172

- OHCHR rule of law tools: 'Monitoring Legal Systems' http://www.unrol.org/doc.aspx?n=RoL+Tools+for+Post+Conflict+States_Monitoring+Legal+Systems.pdf

- DPKO/OHCHR RoL Indicators, '2.1. Source of Data – Filed Data', p.24, and 'Project Tool No. 9 – Field Data Collection' p.117

SECTION 3: USING FINDINGS TO INFORM POLICY AND PROGRAMMING

Political Analysis

"Key to understanding the different approaches to monitoring and evaluation of politically informed programmes is an understanding of the context in which developmental changes take place, alongside an understanding of the purpose and use of monitoring and evaluation."

The Evaluation of Politics and the Politics of Evaluation, 2012, p.8

Measurement is a valuable tool for developing evidence-based RoL policies. However, the process for developing policies can be politicized and is typically influenced by a wide range of factors. Research that is relevant, easily understood and accessible will have the greatest chance of influencing the decisions of policy-makers and practitioners. Producing findings that have clear policy implications will generally not lead to policy change unless policy-makers are engaged and the results cater to their needs and priorities.

Measurement can be used to influence policy and programming in a number of ways:

- **Assessments** might raise problems that require immediate attention, such as, data collection efforts that document challenges with justice institutions' responses to SGBV (see Box 16.) The data collected and shared with national as well as local authorities can inform national actors' decision-making as well as support by international actors.

- **Mid-term Evaluations** can generate information about some of the obstacles to project implementation and generate recommendations for policy changes to ensure project completion. For example, an investigation of a data discrepancy between a government agency and the UN may reveal flaws in data collection by one or more agencies.

- **Final Evaluation** findings can be used to gauge project effectiveness and inform public policy. For example, if, as was the case in South Kivu province of DRC in 2011, an evaluation finds that training judicial monitors to collect data on the judicial response to SGBV in one jurisdiction resulted in an improved data capacity, this may be used to expand similar training initiatives to other jurisdictions. Conversely, if the training is found to be ultimately ineffective, project staff may be able to recommend alternative approaches based on final evaluation results.

Successful and efficient use of measurement findings for policy development and programming requires creativity, substantive knowledge and stakeholder support. Substantive knowledge of an issue will demonstrate project staff expertise, build confidence in project findings, allow staff to discuss policy ideas in a way that practitioners can relate to, and help gauge the validity of suggested recommendations. Finally, support from stakeholders will be essential during the process of developing and implementing policy recommendations.

Box 26, taken from a report by the Overseas Development Institute, describes some of the common pitfalls that can limit the impact of research.

The following consideration can help to plan for policy impact:

> **BOX 26: PEOPLE'S PERSPECTIVE ON ACCESS TO JUSTICE SURVEY IN LAO PDR**
>
> - Inadequate supply of, and access to, relevant information
> - Researchers' poor comprehension of policy process, and unrealistic recommendations
> - Ineffective communication of research
> - Inadequate capacity among policy makers
> - Politicization of research, using it selectively to legitimize decisions
> - Gaps in understanding between researchers, policy makers and public
> - Time lag between dissemination of research and impact on policy
> - Research is deemed unimportant, censored or controlled
> - Some ways of knowing are seen as more valid than others

Take steps to ensure ownership of findings: By involving key stakeholders at all stages of the design, data collection, and analysis phases of measurement activities, the likelihood of findings being taken seriously will be maximized and recommendations acted upon. At a minimum, policy audiences should be aware of measurement activities well in advance of the release of findings. Ideally, they should have an opportunity to review the measurement design, participate in research interviews or other data collection activities, and provide feedback on the findings. Expert interviews provide an opportunity to involve policy audiences by asking them to nominate representatives from their office to participate in interviews (see Section 3.2.C.). The United Nations Rule of Law Indicators Project encourages ownership of project findings by convening 'review panels' including a representative from civil society and a senior government official to determine a rating for 16 of the 135 project indicators that are based on review of legislation and other official documents.

Include information relevant to policy-makers: To maximize the likelihood that recommendations will be implemented it is important to present findings in an accessible format that addresses the concerns of policy-makers. Long, technical research documents are not suited to officials with limited time and multiple competing priorities. Consider supplementing detailed technical documents with short policy briefs. Policy briefs should use non-technical language and include summaries of key recommendations. If possible, include information on the resources required to implement recommendations, the benefits and risks associated with implementation, and projections of the numbers of people that will benefit from the proposed reforms.

Plan for policy impact from the beginning of project measurement: It is important to incorporate policy development into the initial discussions of project and measurement design. Incorporate the concerns of stakeholders, policy-makers and development partners at the beginning of a project to maximize the impact of its findings.

Include measures that are important to policy-makers: Policy-makers will be more receptive to results that are important to their work. This can be as simple as adding an additional item to an observation worksheet or public survey.

Connect measurement aims to existing policy initiatives: It is essential to consider the policy context of measurement activities. By identifying the connections between project findings and the expressed priorities of policy-makers project staff can maximize the likelihood that their recommendations will translate into policy change. For example, a local politician who campaigned for office on an anti-corruption platform may be particularly amenable to adopting recommendations that call for a new independent committee to investigate corporate malfeasance. Wherever possible, findings should also incorporate the priorities identified in an existing poverty reduction strategy plan (or equivalent national development plan) and initiatives to support the Millennium Development Goals.

Coordinate project activities with national and international partners: In most development settings there are multiple international agencies, NGOs and national civil society organizations working on related issues. To the extent possible, it is important to coordinate data collection and dissemination of findings with similarly focused partners. A lack of coordination can easily lead to frustration if senior government officials are contacted by multiple agencies to request the same administrative datasets, to conduct interviews on closely related topics, or to request meetings to discuss duplicative policy recommendations. A coordinated plan, which describes a complementary set of recommendations drawn from multiple agencies and with connections to sources of aid support, will be much more likely to result in policy change.

Describe the experiences of vulnerable and marginalized groups: The results of a nationally representative public survey or series of civil society stakeholder consultations may provide important information on the experiences of ethnic minorities and other under-represented groups. Describing the problems that these groups experience when accessing justice can be a powerful policy lever. These descriptions can be particularly compelling when qualitative, unstructured methods give voice to the perspective of these groups. For example, using the voices of religious minorities displaced by conflict to describe rates of violent abuse can be very impactful.

The policy-making process may be as long as measurement itself and will likely

require a wide range of activities. The table below describes some of these activities, grouped into four broad categories — report development, dissemination of findings, policy recommendation and implementation of policy change. While these categories are not mutually exclusive and do not always proceed in this order (e.g., policy recommendation may precede dissemination), they contain basic guidance on how to trigger meaningful and lasting policy change.

In addition to national level policy change, project measurement can be critical to influencing regional and global level policy development and knowledge leadership,

TABLE 11: SUMMARY STEPS FOR USING MEASUREMENT FINDINGS TOWARDS POLICY-MAKING AND PROGRAMMING	
Category	**Steps**
Report development	- Provide summaries and briefs with a note that the full report can be shared upon request. - Talk about major findings and provide a minimum amount of methodological detail. - If the final evaluation resulted in a wide range of significant findings, think about splitting them into separate reports and possibly for different audiences. - Connect findings to project design and identification of project outputs. - Emphasize both negative and positive findings and frame findings in an encouraging and non-judgmental way. - Provide graphs and other visuals to help understand findings and use non-technical, plain language; do not present findings in English if that is not the primary language in use. - Acknowledge stakeholders, including funders, for their assistance.
Dissemination of findings	- Identify possible audiences for findings, which should influence both content and language of the report. - Determine a medium for the dissemination of findings (oral or written, internal or public). - Introduce the report in-person instead of sending a report via email or post. - Distribute draft reports to stakeholders so they can provide feedback before making the report publicly available. - Discuss findings in a way that is relevant for policy and practice. - Describe challenges of conducting all measurement steps. - Acknowledge limitations instead of defending findings at a later stage.
Policy recommendation	- Convene a roundtable of stakeholders and discuss how measurement findings could be used to influence policy and practice. - Call findings tentative and ensure stakeholder buy-in in the process of finalizing them. - Build consensus across stakeholders, especially around finalized recommendations. - Meet with people affected by a project and elicit their suggestions on how to make a policy change most effective. - Estimate costs associated with proposed solutions. - Determine appropriate time, geographic area and/or institution for triggering change.
Implementation of policy change	- Develop a realistic plan for policy change that reflects resource, political, economic, data and cultural constraints. - Emphasize benefits to an organization as a way to engage its leaders in implementing changes to existing policy and practice. - Provide necessary training and train trainers to maximize impact. - Provide partners with tools and indicators so they can monitor the implementation and impact of recommendations.

as well as UNDP's reporting on its Strategic Plan. There is increasing demand that policy and guidance material be grounded in an evidence base that has emerged from recent programming experience in developing countries. Effective and impactful projects, illustrated by evidence, can provide the basis for scaling-up initiatives and help support South-South and triangular cooperation. They can also be fed into regional policy knowledge generation and experience sharing, as well as the formulation of global UNDP policy and guidance. This will directly help UNDP better communicate its results and its knowledge base and comparative advantage, as well as improve its overall capacity across all regions to implement effective programmes in support of national partner's efforts to strengthen the rule of law.

KEY TAKEAWAYS

In this chapter — How to Measure

• Choose between qualitative and quantitative methods by gauging their strengths and weaknesses
• Measure project effectiveness using baskets of performance indicators to measure broad RoL issues
• Isolate the impact of a project from the effects of other projects
• Distinguish among different types of data
• Conduct administrative data collection, public surveys, expert surveys, focus groups, document reviews and observations

APPENDIX A

SAMPLE TERMS OF REFERENCE

A.1. TOR: ACCESS TO JUSTICE ASSESS-MENT AND BASELINE SURVEY IN THREE REGIONS OF GUINEA BISSAU

1. Introduction

The United Nations Development Assistance Framework for 2008-2012 highlights access to justice as one of the areas that the United Nations Country Team (UNCT) will be working on for the next four years. In particular Output 2.1 of the UNCT Outcome "improved performance of regional and local level structures in fulfilling their role as duty-bearers in delivering services in a transparent and accountable manner", identifies "effective and efficient structures and mechanisms in place and operational to provide access to justice and redress mechanisms" as its target. In order to facilitate the development by government and partners of interventions that address the gaps in the justice system, an assessment is needed to identify the capacities and obstacles of citizens to access the justice system and the capacities of the police, courts, prisons, legal aid services and others to provide justice.

The revised RoLS Programme adopts a people-centred approach that emphasises access to justice in three pilot regions, long-term capacity development and institutionalised training for the judiciary, planning and increased accountability of the sector. It will focus on the formal justice sector mainly through supporting infrastructures and capacity development, but also on the customary and traditional mechanisms, with an emphasis on legal information and access to justice services. Improved governance of the justice and security sector remains as a priority. The policing scope is now limited to the role of the police in the overall functioning of the judicial chain in identified pilot regions. In summary, RoLS prioritize three key areas: i) Decentralization of the justice system and access to justice; ii) Judicial training and mentoring; iii) Strategic planning, coordination and oversight of the justice and security systems.

The access to justice and justice service delivery output of the programme is designed for two principal purposes; to improve the quality and quantity of cases handled by the regional courts, and to improve peoples' access to a remedy for their grievances. The approach under this output aims at the broader sense of access to justice encompassing fair and non-discriminative application of the law; information and civic education about laws and legal procedures; as well as access to the for-

mal justice system and, if preferred, to traditional dispute resolution forums based on restorative justice. This approach is sustained by the regional and international legal framework* as well as the Guinea-Bissauan legislation. The Guinea-Bissauan constitution provides for the right to due process, to legal aid and customary law. Furthermore, specific laws regulate the legal aid scheme. However, in reality poor people lack the resources to claim their rights and the State is not able to fulfil its function and to protect and respect the people's rights.

For this purpose, the output/expected result 1, has a number of sub-outputs, described and summarized below:

1.1 *Capacities of courts and prosecution to administer justice in the regions and communities are strengthened*

1.2 *Legal aid and representation mechanisms, with special focus on women and children, are in place*

1.3 *Alternative dispute resolution mechanisms at community level are strengthened and interface with the formal justice sector*

1.4 *Protection of people's rights and security strengthened through enhanced capacities for law enforcement and criminal investigation*

The administration of justice is a process involving a chain of decisions by several actors. Therefore, the system needs to be addressed as a whole, from the entry point to the end point of the process. Support must be provided to all elements and actors of the process and the linkages between the various actors must been strengthened to ensure a smooth coordination and avoid "bottlenecks" that hamper and slow down the process. UNDP has selected three pilot regions to launch its integrated support to the formal and informal justice sector.

UNDP defines access to justice as "the ability of people, particularly those belonging to poor and disadvantaged groups, to seek and obtain a remedy through formal and informal justice systems, in accordance with human rights principles and standards".

The UNDP Rule of Law and Security Programme will therefore commission a comprehensive regional assessment made up of qualitative and quantitative components. The survey(s) will focus on vulnerable groups, looking at both their capacity to access justice and on the capacities of service providers to deliver justice and will look at the whole justice process from the occurrence of a grievance to the provision of remedies. The findings and the process itself will be used support national partners in the design of national mechanisms for providing equal access to justice.

*Kampala Declaration on Prison conditions in Africa (1996; Dakar Declaration on the Right to a Fair Trial and Legal Assistance in Africa (1999); ACHPR Principles and Guidelines on the Right to a Fair Trial and Legal Assistance in Africa (2001); Ouagadougou Declaration on Accelerating Prison and Penal reform in Africa (2002); Lilongwe Declaration on Accessing Legal Aid in the Criminal Justice System in Africa (2004)

Quantitative and qualitative data generated through the survey will provide a much needed baseline to allow for effective M&E of national and UN/ UNDP interventions in this area. The results of the assessment will also feed directly in to the implementation of the new RoLS Programme for the upcoming two years.

The RoLS Programme is now looking to identify consultants that can support the entire assessment process from start to finish, ensuring coherence across all the steps. In particular, they will work closely will all partners and stakeholders to assist with the design of the qualitative and quantitative parts of the survey, and will take the lead on the data collection and analysis. They will also be responsible for producing the final assessment report which will draw together all assessment findings.

These terms of reference are organised into the following sections. The methodology section outlines clearly the specific steps where the consultants would be required to input.

2.   Objective of Assessment

3.   Methodology

4.   Approach

1.   Team composition, duration and management arrangements

2. Objective of Assessment

The main objective of the assessment is two folded:

1.   To provide for clear baselines in terms of the status on access to justice in the three target regions for monitoring and evaluation purposes

2.   To provide a quantitative and qualitative analysis of the "needs of justice" from the supply side and the demand side of justice in the mentioned regions

PART A: Service Providers

→   To do a mapping (including geographical analysis) of the justice sector in the selected 3 pilot regions: What are the informal and formal institutions in the justice sector from occurrence of grievance to point of remedy.

   ★   Identify dispute resolution mechanisms at the community, sector, regional national levels

   ★   Identify the actors within these institutions and the types of powers they hold

   ★   Identify and map in terms of geographic coverage, community and remote areas that are totally out of range of the formal sector actors and what type of solutions people use for redress of their justice problems

→ To identify and analyze the a) positive factors (what is working) as well as b) obstacles (what is not working) for service providers in both formal and informal systems to fulfill their obligations.

    ★ Identify all strengths and weaknesses of the services providers to deliver the services expected

    ★ Identify awareness, perception and understanding of human rights by the justice sector actors

    ★ Identify accountability mechanisms that prevent abuse of authority by service providers

    ★ Identify incentives and disincentives to ensure responsiveness to those seeking access to justice

PART B: Vulnerable Groups

→ Identify the types of grievances (strictly legal and at what levels, economic, administrative, public administration related etc) faced by the different vulnerable groups (including for example women, female headed households, children and youth, pre-trial detainees, groups disenfranchised due to specific disabilities, extreme poverty and illiterates, displaced persons etc.**)

    ★ Identify the justice perceptions and main priority areas for different groups

    ★ Identify structural problems that contribute to grievances (past conflict, poverty, gender-based discrimination, discrimination based upon ethnicity, denial of citizenship rights, etc.)

→ To identify and analyze the a) positive factors (what is working) as well as b) obstacles (what is not working) for disadvantaged people to access the justice sector to have their grievances redressed

    ★ Identify awareness, perception and understanding of human rights and the justice system by vulnerable groups

    ★ Identify coping mechanisms developed in the absence of recourse to formal justice mechanisms

    ★ Identify sources of conflict that emerge out of the lack of access to justice mechanisms

*** Examples of other criteria for vulnerability (a) The non-fulfilment of basic rights to food, healthcare, education and other government services; (b) Discriminatory treatment of vulnerable and marginalized groups by government or other community frameworks and; (c) The inability to participate in decision-making processes;*

3. Methodology

The assessment will be conducted in an independent and objective manner, always seeking to obtain quantitative as well as qualitative data. It will take place in three regions including Bissau and cover the main sectors of each region (as indicated below). The assessment will follow a human rights based approach—ensuring participation, accountability, equality and non-discrimination. This means that along with conducting participatory consultations and information gathering sessions, the assessment team will also need to share information and hold awareness raising sessions with participants. A continuous process of feedback from these sessions will also need to be channeled to the local and national authorities.

The full process, from award of the contract to the submission of the final report should last no longer than 2 months. Some activities will need to be carefully sequenced, while others can run parallel.

The Research Team is expected to do the mapping of services and target the larger number possible of interviewees and Focus Groups in the following regions, sectors and communities:

Bissau (387,909 population)

- Antula
- Bairro Militar
- Bandim
- Quelele
- Plack I and Plack II
- Misserá

Cacheu (192,508 population)

- Canchungo
- Cacheu
- Bula
- Caió
- Calequisse
- São Domingos
- Bigene

Oio (224,644 population)

- Bissorã

- Mansoa

- Mansabá

- Farim

- Nhacra

Preparatory Work and Tasks

STEP 1. Desk Review of Justice Providers and Key Informant Interviews
(1st field intervention)

1. Identify and establish contacts with the actors in the justice system (formal and informal), including but not limited to:

   a. Groups at the grass-root level (alternate dispute resolution mechanisms), community based organizations and civil society organizations

   b. Local authorities including local government officials and religious and community leaders;

   c. Police force and prosecution;

   d. Prison system;

   e. Court system;

   f. Legal aid providers;

   g. Lawyers, Bar Association;

2. Stocktaking, evaluation of available statistics and preparation of focus groups interviews

3. Primary data/information collection from interviews with justice sector service providers (including prisons, police, courts, legal aid, NGOs, etc.).

STEP 2. Drafting of sample questionnaires and methodology for field research for quantitative and qualitative results

1. Prepare the study methodology and draft questionnaires that will be used for the different focus groups and interviewed and targeting the different objectives.

4. Field Research

STEP 3. Qualitative Data Collection

1. Conduct questionnaires and interviews for service providers (formal and informal) and vulnerable groups

Vulnerable Groups

When brought together, this qualitative research should:

- Identify the range of experiences with the justice sector—formal and informal—as experienced by vulnerable households in Guinea Bissau (assessment of service delivery from the perspective of the vulnerable).
- Gain insights into perceptions about the justice sector—expectations and hindrances—in different situations.
- Explore potential strategies that could be adopted to facilitate the most vulnerable to seek redress from the justice system

The principal method of data collection should be Focus Group Discussions (FGDs), but on particularly complex issues the team may wish to select individuals for more in-depth interviewing. It will be important to ensure that information is collected from both users of the justice system as well as non-users to identify why they have or have not used the system to address their grievances and to identify the obstacles that prevent them from using the justice mechanism.

Key Informants

The team will also need to gather information on how the justice sector works from the perceptions of the service providers. Some of this information may have been collected in Step 1 but most likely there will be a need for open ended interviews with key informants in the legal sector.

To identify the key informants, the research team will draw on their own networks and knowledge of the legal system and also consult with the programme stakeholders.

STEP 4. Finalize Report

Final assessment report to include:

a. Annotated bibliography
b. Assessment Methodology

c.   Mapping of justice sector

d.   Analytical framework

e.   Quantitative findings

f.   Qualitative findings and analysis

g.   Presentation of case studies

STEP 5. National and Regional Dissemination (To be decided during the study imple-mentation)

The findings of the qualitative and quantitative survey will be shared through one national (in Bissau) and two regional workshops with key government, public institutions and civil society stakeholders. Materials will be prepared accordingly.

5. Expected Approach

While conducting the assessment the following should be kept in mind:

- Ensure that both the formal justice mechanisms as well as informal methods of grievance redress are examined (especially in areas where there is no access to the formal justice system and alternative justice mechanisms have been established)

- Be sensitive to the ability/willingness of people to speak freely or even to attend public gatherings

- Ensure adequate safety to those conducting and attending the survey activities

- Ensure objectivity and independence by conducting the assessment in an impartial manner

- Work with local community based organizations to identify participants for the focus group discussions

- In conducting interviews with disadvantaged groups and service providers from the regions be mindful that this is an opportunity to provide information regarding to access to justice and basic legal information and sensitize for the RoLS A2J planned interventions in those regions

6. Assessment Team Composition and Management Arrangements

The research team will be comprised of five elements: one international team leader and four national experts and should be an independent, multi-disciplinary group either affiliated to civil society organizations or to independent bodies, research in-

stitutes or legal companies. The research team will be lead by an international expert consultant on access to justice. The team is expected to have:

- Excellent knowledge of ground situation, justice and administrative systems of Guinea Bissau
- Demonstrated prior experience conducting quantitative and qualitative research
- Demonstrated knowledge and ability to research the access to justice sector and experience of participatory methodologies;
- Experience conducting assessments in partnership with UN and Government;
- Excellent report drafting skills;
- Ability to access the identified regions either independently or through partnership with local organisations.

National members of the team involved in collecting information will go through a workshop training on the concepts and issues around access to justice, the objectives of the assessment, the type of data to be collected, facilitation skills, and sensitivity to gender, conflict and human rights in order to conduct the consultations in a rights-based manner where people are free and comfortable to speak up. The training will also include sensitization and legal awareness skills so that the assessment is also an opportunity for raising awareness to access to justice. This training will be provided at the beginning of the assignment by the team leader and other resources identified within the UN system.

The Research team will work closely with the Ministry of Justice relevant directions and under the guidance and supervision of RoLS Programme Specialist. The Communications and Monitoring Officer of RoLS will also be part of the team and provide backstop support to all extent possible.

The Team Leader will be required to provide regular updates on progress of the works to the RoLS programme specialist.

UNDP will make available to the team all information required in terms of desk review materials and the logistic support. Consultants should have their own laptop.

Terms of Reference for the Access to Justice Assessment Team

Team Leader on Access to Justice Assessments / International Consultant

Duration of contract: 2 months

UNDP, Guinea Bissau

1. Roles, Responsibilities and Outputs

The Team Leader is the principal responsible for the result of the assessment and for the management of the overall team. The main outputs expected are:

1. Development of the assessment methodology based on the guidelines provided on the Terms of Reference
2. Training and management of the Assessment Team
3. Development of the questionnaires and identification and organization of Focus Groups in collaboration with the team
4. Lead all consultation process
5. Final Assessment Report
6. Lead any required actions for dissemination of the main findings

2. Qualifications and Professional Experiences

- Advanced university degree in law with specialization in human rights, international law or social and development studies
- Minimum of 7 years relevant experience in the field of law, specially on access to justice and human rights in developing post conflict countries including training and capacity building on the relevant areas
- Experience in conducting research with multidisciplinary teams on legal and justice issues and in working with government officials and civil society organizations
- Availability and willingness and personal initiative to travel and engage in field work with the community
- Excellent analytical, research, report writing and capacity building skills
- Fluency in Portuguese and in French or English is required.

Terms of Reference for National Consultants for an Access to Justice Assessment

National Consultancy for Legal Experts – 2 posts

Duration of contract: 2 months

1. Roles, Responsibilities and Outputs

As members of the assessment team, the national legal experts are expected to contribute to the overall results of the assessment in particular by bringing in their legal knowledge on the context of Guinea Bissau and their network of relevant contacts in the justice sector critical for the successful implementation of the task. In particular they are expected to:

1. Provide an overview to the team on the functioning of the formal and informal justice system in the country and all critical aspects in terms of information based on their local knowledge and experience

2. Contribute to the development of the assessment methodology based on the guidelines provided on the Terms of Reference

3. Contribute to the development of the questionnaires and identification and organization of Focus Groups in collaboration with the team

4. Assist all processes of data collection quantitative and qualitative

5. Facilitate process of meetings and organization of focus groups in the regions under the guidance of the team leader

6. Facilitate interaction and the successful outcome of the discussions during the meetings and consultations

7. Contribute as guided by the Team Leader to the Final Assessment Report

8. Participate in any required actions for dissemination of the main findings

9. Other tasks oriented by the Team Leader important for the success of the assessment

2. Qualifications and Professional Experiences

- University degree in law. Master an asset.

- Relevant experience in the legal sector as a lawyer or magistrate

- Experience in law teaching or training desirable

- Experience in research and analysis in the fields required and in working with government officials and civil society organizations desirable

- Availability and willingness and personal initiative to travel and engage in field work with the community

- Willingness to learn more and demonstrated interest in human rights, access to justice

- Fluency in Portuguese and Creoule. French or English skills an asset

Terms of Reference for Consultant for an Access to Justice Assessment

National Expert in Social Studies or Development and Economics Policy

Duration of contract: 2 months

1. Roles, Responsibilities and Outputs

As member of the assessment team, the national expert in social studies or development policy is expected to contribute to the overall results of the assessment in particular by bringing in his/hers sociological, political and economic knowledge on the context of Guinea Bissau and the analytical expertise for the successful implementation of the task. In particular the consultant is expected to:

1. Provide an overview to the team on the social and developmental issues in the context of Guinea Bissau relevant for the assessment, in particular the informal and traditional mechanisms related with justice issues

2. Contribute to the development of the assessment methodology based on the guidelines provided on the Terms of Reference

3. Contribute to the development of the questionnaires and identification and organization of Focus Groups in collaboration with the team

4. Assist all processes of data collection quantitative and qualitative

5. Facilitate process of meetings and organization of focus groups in the regions under the guidance of the team leader

6. Facilitate interaction and the successful outcome of the discussions during the meetings and consultations

7. Contribute as guided by the Team Leader to the Final Assessment Report

8. Participate in any required actions for dissemination of the main findings

9. Other tasks oriented by the Team Leader important for the success of the assessment

2. Qualifications and Professional Experiences

• University degree in social, development or economic policies. Master an asset.

• Relevant experience in the field of sociology or economic development

• Prior experience as researcher or professor in social studies desirable

• Experience in working with government officials and civil society organizations desirable in particular related to the justice sector

• Availability and willingness and personal initiative to travel and engage in field work with the community

• Willingness to learn more and demonstrated interest in human rights, access to justice

• Fluency in Portuguese and Creoule. French or English skills an asset

Terms of Reference for Consultant for an Access to Justice Assessment

National Consultant Expert in Human Rights and/or Gender

Duration of contract: 2 months

1. Roles, Responsibilities and Outputs

As member of the assessment team, the national expert in human rights and/or gender is expected to contribute to the overall results of the assessment in particular by bringing in his/hers human rights expertise (in particular women and children's rights) on the context of Guinea Bissau for the successful implementation of the task. In particular the consultant is expected to:

1. Provide an overview to the team and to the assignment on the main and most critical issues related to human rights violations or deprivations, with special emphasis on children and women rights in the context of Guinea Bissau relevant for the assessment

2. Contribute to the development of the assessment methodology based on the guidelines provided on the Terms of Reference

3. Contribute to the development of the questionnaires and identification and organization of Focus Groups in collaboration with the team

4. Assist all processes of data collection quantitative and qualitative

5. Facilitate process of meetings and organization of focus groups in the regions under the guidance of the team leader

6. Facilitate interaction and the successful outcome of the discussions during the meetings and consultations

7. Contribute as guided by the Team Leader to the Final Assessment Report

8. Participate in any required actions for dissemination of the main findings

9. Other tasks oriented by the Team Leader important for the success of the assessment

2. Qualifications and Professional Experiences

- University degree in law or social studies with specialized training in human rights
- Minimum 3 years experience working on human rights issues
- Prior experience as researcher or for the UN or multilateral organizations in the field of human rights and/or gender desirable
- Experience in working with government officials and civil society organizations desirable in particular related to the justice sector
- Availability and willingness and personal initiative to travel and engage in field work with the community
- Fluency in Portuguese and Creoule. French or English skills an asset

A.2. TOR: CONSULTANCY ON ASSISTING DESIGN PEOPLE'S PERSPECTIVE ON ACCESS TO JUSTICE SURVEY IN LAO PDR

I. Background:

In 2005, the Government presented a policy paper on governance indicating its intention to pursue a reform agenda of gradually improving governance through strengthening the Rule of Law. The paper highlighted the need for building efficient, effective, equitable and accessible justice and law enforcement systems. In this respect, emphasis was placed on Government's recognition to improve access to the legal system and decision-making processes, in order to enable all citizens to be able to fully exercise their legal rights and fulfill their legal duties.

Presently, there is some primary information available pertaining to issues confronting the legal sector and access to justice. These include a UNDP assessment of the legal sector (completed in 2003) and the draft Legal Sector Master Plan. However, the bulk of this information has been collected from official documents and reflects the perception or information provided by high ranking officials from Ministries and Departments, Non-Governmental Organizations and the UNDP. Thus far, no systematic and comprehensive study has been undertaken on the people's understanding or perception of their access to justice (including an assessment of access to justice for the poor, vulnerable and disadvantaged people in remote areas including ethnic groups, women and children). In order to redress this, the Lao Bar Association (LBA) plans to implement this year a "People's Perceptions on Access to Justice" (the Survey). The Survey will be used to gather information across all regions of Laos in relation to people's perceptions on the justice system in Laos, including for example:

- their knowledge or understanding of the justice system;
- their interaction with the justice system; and
- their perception of the effectiveness and efficiency of the justice system.

II. Purpose

The LBA does not have experience in designing, planning or implementing access to justice surveys. The purpose of this consultancy is for the consultant to draft a Survey Questionnaire and provide advice and assistance to the LBA in having the Survey planned, designed and implemented, including in relation to:

1. planning for the development through to implementation of the Survey;

2. assessing targets of the Survey;

3. preparation of other Survey Documentation, such as documents outlining the methodology/design for the Survey (including as to size, sample size and other aspects of sampling) and an implementation plan; and

4. assistance in identifying necessary implementing partners to actually conduct the Survey.

The LBA also expects that, as part of the development and implementation process for the Survey, it will be important for appropriate feedback and involvement, and general agreement, to be sought from the Ministry of Justice (MoJ), members of the judiciary, offices of the Public Prosecutor, and other identified stakeholders. The LBA will also require advice and assistance on these matters.

Expected outputs and services for the consultancy are described further below.

III. Outputs and Services

Working closely with the LBA project team and UNDP, the consultant shall provide the following services ensuring at all times close coordination with Customary Law Survey:

- Provide advice assistance on identifying a suitable implementation partner (or the manner in which to advertise a tender, etc.).

- Provide advice and assistance on the terms of reference, tender and/or other relevant documents (technical component) for procurement of the implementation party.

- Draft the Survey Questionnaire (with input from the LBA project team).

- Organize and facilitate an initial consultative workshop in Vientiane to:

    ★ Explain to stakeholders about access to justice surveys;

    ★ Obtain feedback in relation to how the Survey might be conducted and the nature of Survey documentation (using where possible any draft ideas); and

    ★ Thereby, generate awareness and understanding of the Survey and in general agreement with it being undertaken.

Responsibilities for the workshop will include:

- ★ Finalizing the agenda for the workshop, with the approval of LBA;
- ★ Providing input and suggestions for invited stakeholders;
- ★ Facilitating sessions during the workshop (in accordance with the agreed agenda);
- ★ Reporting on the outcome of the workshop.

- Organize and facilitate a second consultative workshop in Vientiane to:
  - ★ Obtain feedback in relation to proposed or drafts of the Survey Documentation, and Survey Questionnaire; and
  - ★ Where possible, obtain buy in for the Survey being carried out.

Responsibilities for the workshop will include:

- ★ Finalizing the agenda for the workshop, with the approval of LBA;
- ★ Providing input and suggestions for invited stakeholders;
- ★ Facilitating sessions during the workshop (in accordance with the agreed agenda);
- ★ Reporting on the outcome of the workshop.

- If considered appropriate, organize and facilitate a third consultative workshop in Vientiane to present the final Survey Documentation and Survey Questionnaire to appropriate stakeholders. If no buy in was possible at the second consultative workshop, it should be sought in this workshop.

Responsibilities for the workshop will include:

- ★ Finalizing the agenda for the workshop, with the approval of LBA;
- ★ Providing input and suggestions for invited stakeholders;
- ★ Facilitating sessions during the workshop (in accordance with the agreed agenda);
- ★ Reporting on the outcome of the workshop.

Working closely with the LBA project team, UNDP and any implementing partner, the consultant shall provide the following services:

- Provide advice and assistance on development of a suitable implementation plan, implementation parties, and design/methodology for the Survey and any other required Survey Documentation. This shall include providing advice in relation to the different types of survey designs and methodologies. It shall also include providing advice in relation to the problems that may be faced by different forms of survey design and methodologies and assisting the LBA to opt between various options.

- Provide advice and assistance on anticipating problems or barriers with implementation of the Survey, in particular, obtaining information from ethnic and regional villages. This may include such things as language, obtaining appropriate sample size, undertaking implementation in a cultural sensitive manner.

- Provide advice and assistance on developing an appropriate budget for implementation of the Survey. This may include providing different options for different types of Survey designs and methodologies.

Working closely with the LBA project team, UNDP and a national consultant, the consultant shall provide the following outputs:

- A plan for the development through to implementation of the Survey.

- Terms of Reference for the implementing party to carry out the Survey.

- The Survey Questionnaire in simple-to-understand English (as it will need to be translated and explainable in the Lao language).

- A final and any necessary interim mission reports, including budget estimations, outcomes and issues arising from the workshops and recommendations. These must be provided in a concise manner.

Assisting the LBA project team to analyze assess and present the results of the Survey.

IV. LBA/UNDP Support provided

The UNDP/LBA Project will provide the consultant with:

- background information on the legal sector in Laos;

- briefings on the other issues in Laos;

- contacts with other organisations that may be of assistance; and

- support in arranging meetings with relevant stakeholders (including workshops).

If considered necessary, and if available, the consultant will also be supported by an anthropologist to ensure cultural context is given sufficient consideration in the survey methodology and subsequent implementation.

V. Qualification and experience

- At least 5 years experience in the legal sector and/or justice related issues.

- Previous involvement in field surveys on access to justice, judicial or legal related work, including relevant experience in the region is essential.

- Experience with various survey methodologies and approaches

- Experience in working with governments and institutions.

- Fluency in spoken and written English including excelling drafting skills, particularly in being able to write documents (such as reports and the questionnaire) concisely and in simple-to-understand English that is easy to translate into another language .

- Understanding of development issues in Lao PDR including cultural and socio-economic environment is an asset.

VI. Duration and date of assignments:

The consultant will commence as soon as possible for a period of an initial two weeks at the beginning of June 2009.

VII. Duty Station:

The consultancy will be located primarily at the LBA office in Vientiane.

VIII. Fee and payment arrangements

The consultancy will be subject to a negotiable lump sum under a UNDP contract entered into on behalf of the project. The fee will be paid in several (at least three) instalments tied to satisfactory completion of specific outputs.

A.3. TOR: PUBLIC PERCEPTIONS OF PALESTINIAN JUSTICE AND SECURITY INSTITUTIONS

I. BACKGROUND

UNDP/PAPP (Programme of Assistance to the Palestinian People) is seeking suitable candidates for the position of International Legal Expert. The position is home-based. Under the guidance and direct supervision of the Chief Technical Advisor/Programme Manager, the International Legal Expert will be responsible for finalizing a survey aimed at measuring public perceptions of key justice and security sector institutions in the occupied Palestinian territory (oPt).

Since 2004, UNDP/PAPP has provided capacity development support to both the High Judicial Council and the Office of the Attorney-General. A UNDP/BCPR mission carried out in 2009 recommended that UNDP/PAPP undertake an expanded programme to fill the gaps in the Palestinian rule of law sector through significant and more substantive changes in the scope, structure and size of its interventions.

Under the UNDP Rule of Law & Access to Justice Programme in the oPt, efforts are directed towards strengthening the institutional capacity of the Ministry of Justice, extending the outreach of legal aid services and legal awareness programmes, setting up the foundation for work with civil society in Gaza Strip, improving gender and juvenile justice conditions and exploring modalities for engagement with the informal justice systems. Assistance is also provided to enhance public confidence in the justice system.

II. SCOPE OF WORK

Under the direct supervision of the Chief Technical Specialist/Programme Manager, and with the support of the M&E Analyst, the International Legal Expert will be responsible for the following:

- Edit and finalise the draft public perception survey, which will be shared with PNA institutions, civil society organizations and academic institutions, and donors/development partners;

- Perform other duties as and when required by the Chief Technical Specialist/ Programme Manager.

III. DELIVERABLES

With the support of the M&E Analyst, the International Legal Expert will produce a high quality report, totalling 40 pages plus annexes, with an executive summary of not more than 5 pages describing key findings and recommendations. The survey should provide, inter alia:

- A clear and concise overview of the data disaggregated by sex, age and location;

- Careful analysis of the data, including concrete recommendations on how to improve public perceptions of key justice and security sector institutions;

- Possible entry points for UNDP/PAPP's Rule of Law & Access to Justice Programme in the oPt; and

- Detailed description of the methodology, including sampling, data collection, data entry, data clearance, etc. This should also include an outline of the limitations of the survey.

IV. TIMEFRAME

25 working days.

V. REQUIREMENTS

Competencies

Core Values and Ethics

- Demonstrates cultural sensitivity and ability to work in a multi-cultural environment
- Supports UNDP's corporate goals and values
- Complies with UNDP rules and regulations and code of conduct
- Demonstrates a high degree of integrity

Teamwork

- Builds effective client relationships and partnerships
- Demonstrates excellent interpersonal skills
- Provides guidance and support to others
- Makes valuable practice contributions to the unit and the office

Communication

- Displays excellent oral and written skills
- Listens actively and responds effectively

Task Management

- Plans, prioritizes and delivers a variety of tasks on time
- Exercises sound judgment
- Develops creative solutions and risk management solutions

Education

- Master's Degree or equivalent in law, political science, social science or in a related field.

Experience

- 4 -5 years of progressive experience in relevant fields, including at least 1 year working on access to justice-related issues;

- Proven record in high quality English writing;

- Demonstrated analytical skills; proven experience in review and analysis of raw data;

- Familiarity with the UN/UNDP system;

- Previous experience working in conflict/post-conflict situations constitutes an advantage;

- Strong understanding of the linkages between access to justice, human rights and human development; and

- Sound knowledge and understanding of the political dynamics in the Middle East.

Language requirements

- Highly skilled in professional English writing. Working knowledge of Arabic constitutes an advantage.

APPENDIX B
TECHNICAL NOTES

B.1. Research Designs for Establishing the 'Causal' Effect and their Feasibility

An evaluation is much stronger if it reflects the views of multiple groups, but this may not always be possible. One-Group Designs (i.e., not having a comparison group) can include only one-time or time-series data collection (i.e., data collection on the same subject more than once). The following table breaks down the practical feasibility of these designs, how they can be used to claim the causal effect of a project, and what level of statistical skills they require.

Design Name	Feasibility (Cost/Ethics/Logistics)	Attribution Certainty (Project caused outcome)	Statistical Skill Requirements
One-group (1 time)	High	None	Low*
One-group (2+ times)	Moderate	Low	Moderate**
Quasi-experimental	Low	Moderate	High***
Experimental	Very Low	High	Low****

*   Typically requires simple percentages
**   Interrupted Time Series can be used
*** Require matching treatment and comparison groups with advanced statistical techniques, or regression analyses for Regression Discontinuity Design
**** Require relatively simple statistical tests, such as T-test for comparison of two groups, or ANOVA for comparison of more than two groups

B.2. Two Common Quasi-Experimental Designs

Design Name	Description
Non-equivalent Control Group Design (NeCGD)	Commonly used design which is similar to Pretest-Posttest Control Group design but has no randomly assigned participants. Two groups are selected and baseline data collected (O1=observation before treatment). Then, treatment is implemented in only one project (X=treatment). Finally, the second observation is conducted and results are compared (O2=observation after treatment). Treatment Group O_1 X O_2 Comparison Group O_1 O_2 If the treatment group has more favourable results than a comparison group, and if these groups are similar, then one can assume that the change was due to the RoL project.
Regression Discontinuity Design (RDD)	Similar to NeCGD but different in a sense that individuals are assigned to treatment and comparison groups using a cut-off score that researchers arbitrarily select based on baseline data collection (O1). For example, to find out if supplying prison inmates with food of a higher nutritional value helps inmates gain weight, consider selecting undernourished inmates using a cut-off score of 50kg. Although these individuals would not be expected to gain more weight than inmates who weighed more than 50kg prior to the project, one would expect them to gain more weight than they would have without it. RDD can determine if this is the case. (Note: this design requires regression analysis, which is a statistical technique for estimating the relationship between two variables while taking into account the effect of other factors, e.g., how the amount of food predicts weight gain while controlling for gender, age, health and other factors.)

APPENDIX C

STRENGTHS AND WEAKNESSES OF

QUALITATIVE RESEARCH METHODS

Research Designs for Establishing the 'Causal' Effect and their Feasibility

Strengths of Qualitative Measurement Methods

Because Qualitative Measurement Methods (QualMM) do not rely on pre-determined 'check-box' style response categories to collect data, they provide an opportunity for research participants to define their experiences, beliefs, attitudes and needs on their own terms. This ability to document participants' individual stories can generate greater insights into the context of a problem and may highlight issues and needs that the project team was not previously aware of. For example, instead of asking respondents whether they think that processes for vetting police officers and other public officials is effective (Yes, No, Maybe) and then reporting percentages, QualMM may describe the way that participants understand "effectiveness" in this context as well as collecting information on ways to improve the process of identifying and removing corrupt and abusive officials from public service. The same applies to a range of RoL concepts — safety, security, protection, trust, fairness, access. In most cases project staff will not have direct experience of the problems that the project addresses; seeking a local perspective on these issues through qualitative interviews can identify needs that may not be known immediately and that may highlight potential obstacles to effective implementation. Conducting an initial round of qualitative interviews as part of the assessment can usefully inform the design of quantitative survey questions, ensuring that large-scale surveys address important concepts and adopt the language and terms used by the study population.

Limitations of Qualitative Measurement Methods

While allowing for an in-depth investigation of complex topics, QualMM also suffer from a number of shortcomings. First, because of their detailed and context-specific nature, qualitative approaches rarely generate findings that can be generalized to other groups of people or settings. For example, it would be unwise to assume that observations of a few trials by mobile courts in one region can apply to other regions. Second, it may be impossible to base predictions on qualitative data collection because of the typically small samples and lack of a random selection of participants. For example, because a project that supports women's rights organi-

Limitations of Qualitative Measures
1. Difficult to generalize to population
2. Inability to predict future events
3. Lack of credibility with stakeholders
4. Extended time of data collection
5. Greater likelihood that personal biases impact evaluation

zations has demonstrated some success in a given time period, it will be difficult to know if the same approach will have similar impacts in the future. Third, while QualMM can be an invaluable source of information on a project's impact, national governments, donors and other key stakeholders are accustomed to thinking about

accountability in terms of quantitative results and may view qualitative findings as less credible. For example, donors may be hesitant to fund a project based on evidence of effectiveness from a small number of interviews with project participants, even if those participants are able to describe the benefits of participation in great detail. Finally, QualMM tend to be time-consuming and researchers collecting qualitative data require specific training on approaches to avoid biasing the results based on their personal views and opinions.

APPENDIX D

STRENGTHS AND WEAKNESSES OF
QUANTITATIVE MEASUREMENT METHODS

Strengths of Quantitative Measurement Methods

Quantitative Measurement Methods (QuantMM) are often the preferred measurement option when changes are tracked over time with a large group of individuals or cases. It is also preferred as a way to produce findings that can be generalized over time, or between groups. QuantMM can be used to reduce complex social phenomena into measurable units to allow for comparison. For example, there may be a need to compare the number of violent crimes in a project site during three consecutive time-periods -— before, during and after a project's operating period.

Example of Quantitative Data

QuantMM can be used to predict future RoL developments and make causal inferences (see Section 3.1.A.
for more on causality). For example, collecting data on robberies and armed assaults over time may help to predict crime rates for subsequent years. Another example could involve measuring how a disarmament project contributes to a decline in violent crime in a region.

Additionally, QuantMM can be fairly quick, especially if there is reliable administrative data provided by other agencies. Because these methods produce data that can be compared over time using standardized measures, they are often considered to be more credible than qualitative information.

Limitations of Quantitative Measurement Methods

However, while QuantMM allow trends to be compared between places and over time, these methods can lose the nuance of the issues being studied. Quantitative measures of violent crime, for example, often include minor assaults and homicides in the same category, conflating two very different problems. QuantMM can also cause problems when researchers do not have a full understanding of the context in which the data are collected. For example, it is essential to understand the complexity and diversity of informal justice mechanisms before developing a survey questionnaire on the use of these systems. A survey finding that "45% of rural residents perceive informal justice mechanisms as responsive to their needs" will mean very little without knowing which systems respondents are referring to. Even the best statistical models and sampling designs will not account for this lack of foundational information. In these cases it is advisable to use more QualMM initially.

QuantMM can also be problematic when using secondary administrative data provided by government agencies or other official institutions. Interpreting this information and understanding potential biases requires a detailed understanding of the data collection and management capacity of these institutions, for what purpose the data was initially collected, and/or competing political interests in measurement results (see section 2.2.D.). Interpreting official statistics can be particularly challenging in conflict-affected and fragile settings where the infrastructure required for collecting and maintaining accurate administrative data may be extremely weak or non-existent. Although some information may be available to police stations, courts and prisons operating in the national capital or other urban areas, records in many countries are not accurately maintained in rural areas. Drawing conclusions based on the results of quantitative data collection typically requires large sample sizes and an appropriate sample selection method, requiring a commitment of resources and advanced methodological skills.

Finally, it is a common fallacy that all types of social phenomena can be quantified. QuantMM can also be too narrowly focused on statistical information without capturing a full picture of peoples' needs, experiences and perceptions.

APPENDIX E

NOTE ON SURVEY METHODS

Measurement can rely on a wide range of methods. These may include observations, interviews, focus groups and reviews of documents, reports or other media. All of these methods can be structured, semi-structured or unstructured.[37]

Structured	Semi-structured	Unstructured
Researchers develop data collection tools in advance. This may include a detailed observation protocol or questionnaire including a series of open-ended questions. Each respondent is asked the same set of questions in the same order.	Some parts of data collection tools and questionnaires are structured but there is some ability to revise or supplement measures during data collection (e.g., based on a review of a few criminal case files, one may revise a strategy for recording data).	No pre-determined protocols or survey tools. The researcher makes decisions about measures during data collection (e.g., how to word a question or arrange a sequence of questions.)

While unstructured approaches tend to be more time-intensive, in terms of both collecting and managing data, they are particularly suited to sensitive issues that require an opportunity for respondents to express their concerns without inhibition. However, because this approach is time consuming and requires experienced researchers, many qualitative data collection tools are structured or semi-structured. Qualitative questions in structured surveys use open-ended questions such as, "Please describe your interaction with the police" to generate a narrative account of respondent's experiences and perceptions. Semi-structured interviews are typically based on a list of questions in a given order that can be modified or re-arranged by the researchers depending on the responses of the interviewee.

Unstructured interviews are often recorded verbatim, using audio recording equipment or through detailed, hand-written notes. Transcripts of audio-recorded interviews maintain the detail of an interviewee's responses and can provide direct quotes that powerfully illustrate the issues being explored in participants' own words. However, transcribing interviews is a detailed and time-consuming process and interviewees may object to being recorded for cultural reasons, or concerns of being interviewed on the record.

---

37 For more detail on structured, semi-structured and unstructured methods, see Annabel Bhamani Kajornboon, 'Using Interviews as Research Instruments', available from http://www.culi.chula.ac.th/e-Journal/bod/Annabel.pdf

APPENDIX F

SAMPLE INFORMED CONSENT FORM FOR PUBLIC AND EXPERT SURVEYS

Informed Consent for [project name]

[Name of the Organization Implementing the Survey]

Interviewer: [Interviewer name]

Date: ___ / ___ / ___

Oral consent was given: Yes _____ No _____

[Oral consent is elicited to ensure the complete confidentiality of a respondent in the absence of a signature]

Signature of interviewer who administered consent: _____

[This is a way to ensure the confidentiality of a respondent who will then not be required to sign his/her name to anything]

Instructions for the interviewer (in italics)
Read the following text to the interviewee:

"[Organization name], in partnership with "[organization name], is conducting interviews as part of [project name].

Interview results will be used to [e.g., inform data collection activities, data analyses, and interpretation].

This interview is [(not)confidential]. We will not record your name or anything that will identify you on the questionnaire. You do not have to answer any of the questions and you may stop the interview at any time. Withdrawal from, or refusal to participate in the study will involve no penalty.

The interview will take approximately [number of minutes].

Do you have any questions or concerns related to participation?" *If yes, answer the question or address the concern raised. You should contact your supervisor if you are not confident in your answer.*

"You can also direct questions that arise in the future to [name of the lead researcher], the principal researcher, at [phone number, email, address]. I will also give you a copy of this form to keep.

Would you like to participate?"

If yes, check "Yes", sign the consent form (see above) and proceed to the interview.

If no, end interview, and say, "Thank you for your time."

APPENDIX F

EXISTING PERFORMANCE MEASURES,

GUIDES AND OTHER RESOURCES

A Users' Guide to Measuring Corruption, UNDP.
http://www.undp.org/content/undp/en/home/librarypage/
democratic-governance/anti-corruption/a-users-guide-to-mea-
suring-corruption/

A Users' Guide to Measuring Local Governance, UNDP.
http://www.undp.org/content/undp/en/home/librarypage/
democratic-governance/local_governance/a-users-guide-to-
measuring-local-governance-/

Bibliography of Governance and Democracy Indicators Resources,
UNDP Oslo Governance Centre.
http://www.lulu.com/us/en/shop/undp-oslo-governance-centre/
bibliography-of-governance-and-democracy-indicators-resourc-
es/ebook/product-18573835.html

Criminal Justice Assessment Toolkit, UNODC.
http://www.unodc.org/unodc/en/justice-and-prison-reform/
Criminal-Justice-Toolkit.html

*Developing and Sustaining Community-Based Participatory Re-
search Partnerships: A Skill-Building Curriculum.*
http://www.cbprcurriculum.info/

Donor Approaches to Governance Assessments, (OECD), 2009.
http://www.oecd.org/dac/governance-development/42472200.
pdf

Governance Indicators: A User's Guide, UNDP.
http://www.undp.org/content/undp/en/home/librarypage/dem-
ocratic-governance/oslo_governance_centre/governance_as-
sessments/governance-indicators-2nd-edition.html

Kellogg Foundation Logic Model Development Guide.
http://www.wkkf.org/knowledge-center/resources/2006/02/WK-
Kellogg-Foundation-Logic-Model-Development-Guide.aspx

Manual for the Measurement of Juvenile Justice Indicators, UNODC/
UNICEF, 2007.
http://www.unicef.org/albania/JJ_Indicators_manual.pdf

*Measuring Progress toward Safety and Justice: A Global Guide to
the Design of Performance Indicators across the Justice Sector, Vera
Institute of Justice,* 2003.
http://www.vera.org/pubs/measuring-progress-toward-safety-
and-justice-global-guide-design-performance-indicators-across

Methodological Guidelines for Local Governance Analysis, UNDP,
2007.
http://gaportal.org/resources/detail/methodological-guidelines-
for-local-governance-analysis

Monitoring and Evaluation when Politics Matters, by Chris Roche
and Linda Kelly, Developmental Leadership Program, 2012.
http://www.dlprog.org/news-events/monitoring-and-evalua-
tion-when-politics-matters-notes-from-program-experience.php

*Programming for Justice: Access for All: A Practitioner's Guide to a
Human Rights-based Approach to Access to Justice,* UNDP.
http://regionalcentrebangkok.undp.or.th/practices/governance/
a2j/tools/index1.html

*Result-based Management Handbook, United Nations Develop-
ment Group,* 2011.
http://www.undg.org/content/programming_reference_guide_
(undaf)/un_country_programming_principles/results-based_
management

*The Evaluation of Politics and the Politics of Evaluation, by Chris
Roche and Linda Kelly. Developmental Leadership Program,* 2012.
http://www.dlprog.org/ftp/view/Public Folder/The Evaluation of
Politics and the Politics of Evaluation.pdf

Toolkits on Diversion and Alternatives to Detention, UNICEF.
http://www.unicef.org/tdad/index_55653.html

*United Nations Rule of Law Indicators; Implementation Guide and
Project Tools,* DPKO and OHCHR.
http://www.un.org/en/peacekeeping/publications/un_rule_of_
law_indicators.pdf

Users' Guide on Measuring Fragility, UNDP and DIE.
http://www.undp.org/content/undp/en/home/librarypage/
democratic-governance/oslo_governance_centre/governance_
assessments/a_users_guide_tofragility.html

Want Influence? Eliminate Your Blind Spots, by Kristen Grimm.
Spitfire Strategies, 2013.
http://www.spitfirestrategies.com/influence.html

Empowered lives.
Resilient nations.