43707

Justice Sector Assessments

A Handbook

Dory Reiling
Senior Counsel
Justice Reform Group
Legal Vice Presidency
The World Bank

Linn Hammergren
Senior Public Sector Management Specialist
Latin America and Caribbean Region
The World Bank

Adrian Di Giovanni
Counsel/Legal Associate
Operations Policy Group
Legal Vice Presidency
The World Bank

CONTENTS

Boxes

Box 1: Main steps in assessment
Box 2: Be aware of the impact of who performs the assessment
Box 3: Assessment fatigue
Box 4: Patience sometimes pays off
Box 5: United Nations assessment of Nigerian justice: Survey-based approach
Box 6: Reporting Challenges
Box 7: Appendixes that should be included in a Justice Sector Assessment report

Figures

Figure 1: Growth in caseload in Brazilian courts: First-instance filings, 1941–2001
Figure 2: Average caseload per judge across Brazilian court systems, 1999–2003
Figure 3: Mexico debt collection cases: Formal and informal exits in three moments of the proceedings

Tables

Table 1: Factors that impact the assessment process

Table 2: Guidance on various assessment models
Table 3: Long pretrial custody by country, 2005

Reforming justice sector institutions is considered essential for the World Bank's mission of poverty reduction and economic growth. Nevertheless, the experience of many countries around the world shows that actual reform is not always easy and straightforward. A thorough assessment can provide understanding, a basis for reform programming and consensus building, and a baseline for measuring reform progress.

Since 1994, the World Bank has performed many justice sector assessments, usually on an ad hoc basis and without a common framework. After 12 years, the existing assessment practices were in need of examination. It also was time to identify lessons learned. This handbook reflects justice sector assessments as they have evolved in the World Bank and other institutions and provides a framework for good practice.

Most of the Bank's justice sector assessments have focused on the judiciary and on justice ministries. Experience with other justice sector institutions, like the prosecution, and with traditional, informal, non-state forms of dispute resolution is still limited. The Bank's Legal Empowerment for the Poor initiative and more involvement in criminal justice may introduce new needs. The methodology in this handbook will be helpful for these needs as well.

This handbook is a practical guide, primarily intended for World Bank staff involved in justice sector assessments. It also may be of interest to the wider justice reform and development community. Assessment methodologies for other sectors and justice sector assessment methodologies from other institutions have informed the handbook. As far as we could ascertain, this is the first time that practices in justice sector assessments as they have evolved have actually been described. This handbook is not the last word in assessments; rather, it is a basis for further development.

Olga Ruda, Justice Reform Index Coordinator at the American Bar Association (ABA) Rule of Law Initiative and Jean-Jacques Dethier (DECRS), reviewed the draft. We are much indebted to them for their insightful, incisive, and helpful comments. They caused major changes in the draft. In an early phase, Hiram Chodosh, presently Dean and Professor of Law, S.J. Quinney College of Law, University of Utah, provided much interesting and useful material on justice sector workings and diagnostics. We have used it as input for this handbook. We also thank Richard Messick (PRMPS) and Adolfo Rouillon (LEGPS) for their comments.

Writing this handbook would not have been possible without the input from all members of the LEGJR Practice Group and the many others involved in justice reform in the World Bank who participated in workshops and commented on early drafts. We thank all of them, but particularly Randi Ryterman (PRMPS) and David Bernstein (ECSPE), for sharing their thoughts and experience.

Dory Reiling (LEGJR), Linn Hammergren (LCSPS), Adrian Di Giovanni (LEGOP)

ACRONYMS

ABA	American Bar Association
CAS	country assistance strategy
CAT	Convention Against Torture
CBO	community-based organization
CCJE	Conseil Consultatif de Juges Européens (Consultative Council of European Judges)
CEDAW	Convention on the Elimination of All Forms of Discrimination Against Women
CEJA	Centro de Estudios de Justicia de las Américas (Latin American Center for Justice Studies)
CEPEJ	Commission Européenne pour l'Efficacité de la Justice (European Commission for the Efficiency of Justice)
CMU	Country Management Unit
FATF	Financial Action Task Force on Money Laundering (OECD)
FSAP	Financial Sector Assessment Program
ICCPR	International Covenant on Civil and Political Rights
ILANUD	United Nations Latin American Institute for the Prevention of Crime and the Treatment of Offenders
NCSC	National Center for State Courts
OECD	Organisation for Economic Co-operation and Development
TTL	task team leader
UNODC	United Nations Organization on Drugs and Crime

1. INTRODUCTION

Why This Handbook
The World Bank carried out its first assessment of justice sector institutions in 1994. Justice sector assessments sometimes are used in the Bank in developing a Country Assistance Strategy (CAS), which provides a framework for Bank action in a country. Ideally, an assessment constitutes the first step in a process toward a program of reforms in the justice sector. Since 1994, many justice sector assessments have been done, usually ad hoc and without a common framework. Now that the Bank and other organizations have gained considerable experience in carrying out justice sector assessments, it is timely to examine existing assessment practices and the inventory methodologies used in evaluating justice sector institutions and institutions of other sectors; and to identify lessons learned.

This handbook is the result of that process. It reflects justice sector assessments as they have developed in the World Bank and other institutions and provides a framework for good practice. As far as could be ascertained, this is the first time that practices in justice sector assessments as they have evolved have been described. If some of the advice seems basic and elementary, this may well be the reason. This handbook is conceived as both a stock-taking and a basis for further improvement.

The Justice Sector
"Justice sector" is the term that economists use for what social scientists usually call the justice system. For this book, we define the justice sector or system most broadly as:

> *The institutions that are central to resolving conflicts arising over alleged violations or different interpretations of the rules that societies create to govern members' behavior; and that, as a consequence, are central to strengthening the normative framework (laws and rules) that shapes public and private actions.*

This is the definition most commonly used in the World Bank. Justice sectors or systems are conglomerates of institutions, produced by long historical processes, mostly in a national context, with structures that are unique reflections of the national and historical context. There are some tasks that justice institutions in every nation carry out, but there are great differences across countries as well. Legal systems in Europe and in some civil law countries in Africa, Asia, and Latin America share common frameworks, as do some common law systems. However, national practices have developed that are very different from one another. That there is no one ideal, standard justice system considerably complicates the assessment challenge as it eliminates the possibility of evaluating structure and function against a universal yardstick.

This handbook provides primarily examples from the judiciary because that is the institution on which most of the Bank's past assessments have focused. There is not

much experience yet with other justice sector institutions like the prosecution. However, the guidance in this handbook applies to the whole of the justice sector and all the institutions that belong to it in a given national context.

Assessing the Justice Sector

Assessing justice sectors is a form of social, rather than legal, research. Its subject is the workings of institutions as organizations. The assessments' subject has consequences for its methodologies. Assessing justice sectors is also an emerging field. Consequently, the methodologies are still very much in development. The same holds for any standards that may be used in the assessment. The World Bank's Financial Sector Assessment Program (FSAP) uses an indicative set of questionnaires for some parts of its assessments (World Bank 2005c). Anti-money-laundering assessments investigate the extent of implementation of the 40 recommendations of the OECD's Financial Action Task Force, known as the FATF 40 (FATF 2006). However, justice sector assessments are a different matter. They have no checklist of standards to be ticked off because such rigorous standards do not exist.

No Ideal Model

Having an ideal model would make the assessment process far easier and lend itself to the creation of a single assessment template. However, there is no consensus, in the global justice community or among justice scholars, on what this model should look like. Even in an organization such as the World Bank, the views of those engaged in justice sector work differ markedly.[1] The Bank's implicit model is in reality made up of several different models; the models occasionally seem to work at cross-purposes, as when special courts for commercial cases offer enhanced access to elites or when successful litigation to access second- and third-generation rights threatens to bankrupt the public treasury. The values commonly pursued in any justice reform program - access, efficiency, fair and high quality judgments - are themselves often in conflict and therefore unlikely to be advanced equally over the short to medium run. Over the longer run, perhaps all good things come together, but reform programs do not work in that timeframe.

Issues and Outcomes

Looking beyond system outputs and outcomes to their impacts on broader societal goals such as economic growth, political stability, and poverty reduction, most experts now admit a measure of uncertainty. Even something so apparently obvious as the criminal justice system's impact on crime rates increasingly is questioned by criminal justice experts themselves.[2] It may be fairly easy to identify specific areas in which systemic

---

[1] See for example Santos (2006) for a discussion of the Bank's disparate approaches to the sector's role and reform needs.

[2] : "No one, so far as I am aware, has ever explained a nation's crime rate by taking systematically into account all of the variables that might affect it…. We are getting better at evaluating such efforts and

dysfunctionalities impede realization of these goals. However, identifying the dysfunctionalities is not the same as stipulating that a holistic justice reform is the best way to advance the broader societal goals. This and the lack of broader certainty about the sector's universal contributions thus argue for a problem-focused approach drawing on the complaints most commonly voiced by the citizens of a specific country. There may of course be problems that the World Bank considers outside of its mandate or solutions that the Bank will not support because they are technically unsound or contrary to its own operating principles. As is all reform, justice reform is, in the end, political because it inevitably produces winners and losers. Thus, as do all external cooperators, the Bank needs its own standards. However, in the absence of a universal model, the Bank will be best guided by working in areas identified by a majority of citizens of a particular country as problematic.

Basic Assumptions
This handbook starts from the assumption that a well-performing justice sector, including the judiciary, has the same basic requirements as any organization. It needs to select the correct human resources, manage them for performance, have adequate administrative systems and resources, and organize its fundamental business and operating procedures to achieve the desired outcome.

There does exist a common international legal framework that provides general standards, that is, some guidance as to what justice systems should be able to do. The international human rights conventions and other, more specific instruments contain some norms for justice systems.[3] Basically, what the international legal standards require is quite similar to the basic requirements discussed earlier. Both of these "standards," basic as well as legal, are phrased as abstract ideals: *adequate human resources management, justice with reasonable promptness.* These standards will need to be operationalized to a level at which they provide a framework for looking at reality in those terms: Are there policies and institutions to safeguard those ideals? Are the policies and institutions implemented and used? Are they maintained and sustained?

About Disputes
Whether dispute resolution is effective in that it meets the needs of the users in society is another, albeit related, matter. Most disputes do not come to court. The justice sector and its institutions, individually and as a whole, are part of a larger entity. Research into dispute resolution in developing countries, for instance by the World Bank's Justice for the Poor Program, (World Bank 2006) is deepening the understanding of the need for dispute resolution; and the understanding of traditional, informal, customary dispute resolution systems.[4] In the U.K. and the Netherlands, less than 4 percent of all problems that could give rise to a civil, commercial, or administrative court case actually end up in court (Genn 1999, Van Velthoven 2004). Of the cases that do come to court, fewer than 5 percent are decided by some form of examination of evidence or witness hearing.

---

drawing lessons from them, but we have no idea what would be the cumulative effect of putting in place all of the best programs on a large scale." Wilson (2004), 546–47.
[3] A list of the most important ones for the justice sector can be found in appendix 2.
[4] Systems other than any national, unitary, formal legal system.

Furthermore, the court's decision resolves the dispute in only half of the cases brought to court.

More on the Handbook
This handbook is a practical guide intended primarily for those who are actually involved in justice sector assessments in the World Bank and elsewhere, whether they are carrying out the assessment themselves, or contracting for it. The handbook also may be helpful and informative for others who need to know about assessments or about aspects of the justice sector.

There is more than one possible approach to analyze a justice sector. One might seek to examine the way in which all of the institutions are performing in the system. Alternately, one might identify a particular development problem, such as corruption or land tenure, and analyze the way in which the justice system is performing with respect to it. Most of the guidance in this handbook applies to both approaches.

Determining what the problem is can be as difficult as identifying its causes or devising remedies. The handbook provides detailed advice on methods and processes for the assessment process that help to acquire a better understanding of what is wrong with a justice system and why. The book explains, for example, causality and incentives. It also discusses ways to verify perceived problems and proposed solutions, and provides an overview of methodological tools and research methods.

The primary audience for this handbook is Bank staff who are carrying out a justice sector assessment, or TTLs who are contracting for an assessment. Although some team members may have participated in other assessments, those with sufficient experience to be considered "assessment experts" are still quite rare. Furthermore, having gained their expertise on the ground, they frequently have settled into their own routines without considering the full range of options. Usually chosen for their knowledge of some aspect of judicial operations, many team members may be relative assessment novices, and all might benefit from a greater focus on the collective rather than their individual products. This said, the handbook does not purport to provide an assessment template, but rather to offer general guidelines and lessons of experience to enhance the utility of the team's work.

This handbook follows the order in which the assessment events will take place:

- *Chapter 2* provides considerable guidance on the preassessment stages: choosing a team, setting the assessment scope, organizing, making first choices, managing risks, and building relationships with counterparts. This chapter will be of special interest to the primary audience. *Chapter 3* discusses the assessment process and its different steps in detail. These steps, as well as the core material on the assessment process, also should be useful to the members of the assessment team itself.
- *Chapter 4* examines methodologies and tools and their strengths and weaknesses.

- *Chapter 5* provides concise information on the most frequently encountered issues in justice sector assessments: case delay, access to justice, and corruption.
- *Chapter 6* offers very practical guidance, intended primarily for team leaders, on the post-assessment process of discussion, broader dissemination, and operationalization of the recommendations.

Change and Reform

If the purpose of a reform program or project is to improve sector performance, that of the assessment is to provide a better understanding of what is wrong and why. An assessment is both process and product. The process investigates and analyzes the condition, situation, or problems in, in our case, a justice system. Its product is a picture of a system. It is limited in time and scope. After laws change or reforms have been implemented, the system will begin to change. The assessment will be a basis for tracking this and other changes as well as a baseline against which to measure progress in a subsequent project. The scope of the assessment will be determined largely by what is known before the start and by what is brought up for examination. Over time, this perspective may shift. An assessment thus can play an important role in keeping the change and reform process on track and in evaluating the impact of a reform.

2. GETTING STARTED: PRELIMINARY CONSIDERATIONS

The aim of this chapter is to help the user begin *planning* the assessment. Early attention to the assessment's possible risks and uncertainties will make the process run more smoothly and save time and resources. This chapter provides guidance on constructing a work plan by (1) describing the main stages of the assessment and (2) highlighting the tasks and considerations that should be addressed at each stage during the planning. process.

In this case, an assessment investigates and analyzes the condition, situation, or problems in a justice system. The assessment is an iterative process: information gathered to answer questions will uncover unforeseen results and issues that will give rise to new questions. This may even lead to a shift in the focus, or alter components, of the assessment as it progresses. Spelling out the objectives and scope of the assessment and planning the assessment's other phases from the outset will help to guide choices, providing for a margin of maneuver while keeping the assessment on track.

Box 1. Main steps in assessment

Defining objectives scope
- Building assessment team/expertise
- Identifying stakeholders
- Determining assessment design and methodology
- Performing research
- Report writing
- Dissemination

Among all of the considerations influencing the design of the assessment, the need to *generate support for the assessment* stands out. It will have to be taken into consideration both at the outset and at each stage of the assessment. Other general considerations will each have a direct impact on defining the objective, scope, and focus of the assessment. They are included in Table 1.

Table 1. Factors with an impact on the assessment process

Reform environment	*Assessment process*
▪ Level of development	▪ Any initial instructions from the funders or counterparts on content or methods[6]
▪ Other urgent priorities (political stability, curbing violence, ending extreme inequity)[5]	▪ Time
▪ Existing reforms and reform actors	▪ Budget
▪ Receptivity or fatigue for reform or assessments	▪ Availability of local resources, support, or participants
▪ Local capacity	▪ Team composition
▪ Available donor and local resources	▪ Logistical obstacles: size of country, areas that cannot be visited for security reasons

1. Defining Objectives, Scope, and Problem-Set of the Assessment

The challenge in defining the objectives of an assessment is to be *realistic*, in other words, to balance what would be the ideal outcome with what can realistically can be hoped to be accomplished in the given context within the time, resource, and other constraints of the assessment. The ultimate role of assessments is to guide the reform process and feed into discussions over how to prioritize and direct reform efforts. The *general* objectives of assessments include developing a knowledge base to:

   a. Identify or verify perceived areas in need of reform or deeper analysis
   b. Tailor and sequence reforms more effectively
   c. Establish an empirical baseline against which the outcomes of reform interventions can be measured.

Beyond these general types of objectives, the objectives of an assessment will be very much *case specific*. In all instances, a main concern in setting more specific yet realistic objectives will be deciding how general or narrow the focus of the assessment should be and how deeply to go in performing the research. A main choice is between whether (a) to attempt to evaluate the justice sector as a whole, or (b) to narrow the focus to a narrower and discrete set of issues or problems.

Choosing the *focus* or *topic* of an assessment is a bit of a "catch twenty-two." The general aim of assessments is to identify and analyze problems for reform. How then to choose the topic without having already performed the analysis? As a basic guide: a more general focus would attempt to cover a wider set of issues, institutions, and actors. The risk with this type of broader focus is a more superficial evaluation.[7] In a more specific focus, it is easier to deepen the enquiry. The risk in this case is that the assessment will

---

[5] These additional priorities may either reduce the attention and resources for justice reform or set the goals that it will be asked to address.

[6] Aside from instructions on what is to be covered, funders often insist on certain methodologies (a survey, workshops, use of local NGOs or research groups). They also may have instructions as to what should not be covered, due to political sensitivities or other reasons.

[7] The World Bank's Legal Modernization Initiative notes that the "lack of focus on a particular problem or problems may explain disappointing outcomes." The initiative suggests a shift to a more tailored, problem-centered approach. Bank-Wide Legal Modernization Initiative Justice Sector Action Plan, 8 (Draft, March 8, 2005).

miss key problems or causes outside the narrower focus. In some respects, the distinction between broad and narrow is a false one. As the research goes deeper, a focus on the justice system as a whole likely will lead the assessment to focus on a few priority issues. A more specific focus will likely lead, at some point, to an analysis of how a specific topic or set of issues relates to and is shaped by the system as a whole. Chapter 3 will provide detailed guidance on how to choose.

Table 2. Guidance on various assessment models

Model	Time & cost of study	Pros/cons	Practical tips
#1 Super-deluxe, bells and whistles	6–12 months collecting data (additional time always to produce report) Ample budget ($200,000 +)	Very rich research, broad scope, and potential to go very deep. Opportunity to build/enhance in-country data collecting and survey capacity. Information may be out of date by the time it is reported. May lose interest of stakeholders or opportunities to use report as a vehicle for generating interest in reform.	Have individual members present interim findings and reports to the rest of the team so that collective decisions can be made on areas to explore further. Keep the team on track via an initial plan and set of objectives. This structure can discourage members going off on tangents that, in the end, may not be useful. Provide interim reports to outside stakeholders to get feedback and maintain interest.
#2 Bare-bones, emaciated	2–4 weeks collecting data (additional time always to produce report) Very small budget ($30,000–$50,000)	Problem or situation examined may only require a quick and more limited investigation. May be the only opportunity provided for doing an assessment. Useful as preliminary assessment to identify existing sources of data and possible further stages of diagnosis. Drawbacks: Results may be overly superficial and not rich enough to enable recommendations or action based on findings.	Be upfront about limitations of coverage in the report and highlight need for and map areas of potential follow-up research. Start with a solid plan and avoid diverging from it. When time and resources are scarce, tangential investigations are more risky. Local expertise is always important, but arguably more essential here because of the time it can save. However, where local experts are not versed in assessment methodologies, they may have to be given additional guidance or training. Although likely more cost-effective, local expertise should not be used just to cut cost, but primarily to add value and build capacity in the country.

Desk Reviews

A desk review or stocktaking of existing written and electronic material relevant to the assessment should be performed as an initial step. The desk review will feed into all components of the assessment and planning for it. The review should, where relevant, take stock of: country conditions; the structure and history of the justice system; and existing reports and assessments on the justice system, existing projects, reform efforts

and actors. One main reason for performing desk reviews is to assist in refining the scope and topic of the assessment. Albeit a painstaking task, the review will provide an idea of which assessments and reforms already have been performed or attempted. This knowledge can help to avoid duplicating existing work, to identify stakeholders and other donor actors, and to identify the types of questions and data to target.

Within the last several years, much has been written about particular legal systems and their problems. In some cases, background information about judicial and legal systems is readily available.[8] In addition to the literature on legal and judicial systems and studies from both international and domestic sources, numerous online resources exist. Many countries now maintain their own official websites on the basics of their legal systems.[9] Increasingly, there also are websites devoted to compiling regional statistics and basic overviews.[10]

The review should take stock of existing *legislation* that has a bearing on the assessment's topic. Obvious starting points include constitutional provisions and any law regulating judicial services. Local counterparts are the best source to identify additional pieces of legislation that have a less obvious connection to the assessment's focus. The main reason for reviewing legislation is that it can help to clarify why a current state of affairs exists, for example, bad legislation might be hindering sector performance. The review also guides thinking, from the beginning, of what reforms could more practically or realistically be achieved and thus recommended in light of the later findings of the assessment. Anchoring diagnostics in pre-existing law, for example, guarantees of judicial independence or criminalization of bribes, also can be a source of authority and legitimacy for the assessment. Legislative provisions can serve as benchmarks against which to evaluate conditions; that is, compliance with or fulfilment of legal standards could serve as one mark of the performance of the justice system. However, relying on legislation should be done with extreme care because prevailing conditions quite often are shaped by extra-legal considerations and external circumstances. In many contexts, the laws may be too far removed from real practice to serve as a useful guide, in which case their value is at best as an official expression of aspirations of how things should be.

Desk reviews have *limits*. First, many relevant documents are not easily available. More in-depth studies or those containing empirical data might be accessible only locally through personal contacts and with persistence. Second, preexisting assessments and other documents can be hard to use or unreliable. There may be no way to verify the quality of these documents; their data and conclusions may be inaccurate, out of date, and hard to evaluate. That said, existing documentation on the systems that will be assessed

---

[8] For example, Herbert Kritzer's collection, *Legal Systems of the World: A Political, Social, and Cultural Encyclopedia,* offers a good starting point.

[9] If an analyst is looking for a judicial opinion to evaluate from Papua New Guinea, s/he can find it at www.paclii.org/pg/cases/PGSC/2005/1.html. If a consultant wants to learn about Bhutan's legal system, s/he might click on www.country-data.com/cgi-bin/query/r-1521. html

[10] For example, CEPEJ features information on 46 European countries. CEJA is producing annual reports on the situation in Latin America and the Caribbean.

provides the team with useful starting points. The team should keep in mind that the review stage is too early for conclusions[11].

Building Assessment Team and Expertise

Assessment teams ideally should be *multidisciplinary* and draw on a good mix of expertise and skills. Ideally, all team members should be able to function in the local language, although in some cases this may simply be impossible. At a minimum, some or all team members should possess an understanding of the legal system or the substantive areas of focus in the assessment. Research and social science skills are crucial to ensure the integrity of the assessment results. Depending on the focus, it might be useful to have specialists who have a more general foundation in institutional reform or human resources management; governance and anticorruption; social work; anthropology; and human rights. Again, the composition of the team will depend to a large extent on the focus and depth intended for the assessment.

Local expertise offers many advantages, not the least of which is an ability to work in the local language(s) and an understanding of terms and concepts that could be lost in translation. Local experts can provide more nuanced insight drawn from on-the-ground experience with realities and problems. Depending on the context, employing local experts will build local research capacity and exposure to issues while being more cost effective. Where needed, an initial stage of an assessment might provide a training session on survey techniques to team members. On the one hand, drawing on local expertise can increase the perceived legitimacy and reception of the assessment by not making it appear to be imposed from the outside. On the other hand, a possible tension is that local experts also might have a stake in the assessment's outcomes, in which they or their organization are already players in the reform process and internal politics of the justice system. This could undermine the assessment's perceived legitimacy. The initial identification and analysis of stakeholders can help guide the choice of the appropriate local experts. Where a country expert is unavailable, an expert in a similar system can save a great deal of time in getting up to speed.

International or *comparative expertise* is also valuable: it provides a broad perspective combined with an appreciation of multiple justice systems and reform experiences. International experts also can bring skills and outlooks that are not available in the assessed country, such as court administration, information and communications technology (ICT) applications to judicial operations, or the construction and interpretation of judicial statistical systems.

Successful teams share some common factors apart from their technical expertise:

- Shared understanding of the assessment's purpose, as reflected in their terms of reference
- Agreed set of norms and ground rules

---

[11] "A User's Guide to Poverty and Social Impact Analysis" (World Bank 2003c, 16–17).

- Commitment of all members to the collective task and a willingness to put it before any individual agendas
- Coverage of all necessary roles in the team.

The primary role to be fulfilled is that of moving the team to accomplish its task. The environment and the resources that help the team get its work done must be assured. This is the role of the team leader, who also keeps the team on track. Since the team leader is ultimately responsible for the product, when the team fails to reach agreement by consensus, the team leader ultimately has the last vote. However, team leaders are well advised to never "pull rank" or apply undue influence on the team.

A role that tends to be overlooked is that of *process facilitation*. In small teams, the roles of both leading and facilitating will be done by the team leader. However, the role of facilitator is distinct from that of leader. While the leader's responsibility is the team's result, the facilitator's primary focus is the team's process. The facilitator keeps all team members on track, makes things happen with ease, and helps the group with the process, for instance by moderating discussions. The role of facilitator also may rotate from member to member, thus enabling the team leader to participate fully in the substantive discussions. In large teams with complex processes, the facilitator sometimes is not a team member but is contracted specifically to perform this role to facilitate the process.

Another necessary role in the team is that of *record-keeping*. Notes, documents, and decisions must be kept. The team's key points, ideas, and decisions must be written down and kept for future reference. The team's process, discussions, and decisions need to be documented and kept. Thorough, accurate, and well-organized record-keeping will help produce a good report.

If funding permits, having a *research assistant* for the desk review and possibly later research tasks can be a great help. The role of the research assistant is discussed in more detail in Chapter 3.

Although not absolutely essential to the team's success, other roles also can be very helpful, In a busy team meeting in which members are considering ideas, brainstorming, and prioritizing, time-keeping is very helpful. Teams can be helped with other, more informal roles, too. These include experts who can clarify technical issues, a "big picture" person who helps to summarize items, people who are good at keeping a group together, and experienced negotiators who can help bring opposing views together. If there is a choice among candidate team members, taking such skills into account in the choice may enrich the team and, by adding to its diversity, make for greater success in performing the assessment.

> **Box 2. The impact of who performs the assessment**
>
> Who does the assessment obviously has enormous consequences for its contents. Unfortunately, team membership usually must be decided before the needs are fully understood. If the problem is that a critical expertise was not included because it was identified after the assessment began, the reasonable alternatives (getting the other members to try to fill in or adding another member) may not be feasible. The "hole" will remain identified but not filled. Team members

also can insert biases that, if they are in fairly specialized areas, others may not notice. An example is the software, bankruptcy, or human resources expert who believes that there is only one reasonable way do things, regardless of costs, local capabilities, or incompatibility with national practices.

The composition of the team also may affect how the assessment is received politically—especially where there is a perception that team members have a vested interest in the outcome of the assessment. Careful selection of members, assurance of a reasonably wide representation of skills and experience, and review of their prior work is always advisable. Nevertheless, if mistakes are made, the team leader will need a variety of skills to overcome them.

Identifying Stakeholders and Political Support

From the beginning of the assessment, the involvement of key stakeholders, the local legal community in particular, will help to generate and maximize political support for the assessment. For instance, support will be needed when approaching people to provide input, for example, as interviewees or making administrative data available; for the reception of the assessment results at publication; and for subsequent reforms based on the assessment. This is particularly the case when the assessment is performed with the longer-term intent of improving institutional performance. Thus, from the beginning, it is important to reach out to and involve a broad range of constituents, including ministry officials, court personnel, legal professionals, nongovernmental organizations, advocates of underrepresented groups, business interests, users of the courts, and the press.

There are number of basic ways by which people have a "stake" or are stakeholders in the assessment, such as being:

- Part of the community of practice affected by the issues being assessed
- Interviewees or providers of input to assessment
- Members of the assessment team (discussed above)
- Subjects of the assessment (viz. being assessed)
- Recipients of eventual recommendations or actors to bring forward the recommendations
- Funders of the exercise.

The following are some examples of stakeholders in an assessment:

- Ministry of Justice officials or members of judicial councils, where one exists
- System-users such as plaintiffs and defendants, and the organizations that represent them
- Legal aid providers: lawyers, paralegals, community-based organizations (CBOs)
- Prosecutors
- Leadership and members of the judiciary
- Court management and staff
- Providers of non-court dispute resolution
- Legal training providers
- Individuals with legal issues not reached by the justice system

- Other donors in the justice field.

Box 3. Assessment fatigue

Care needs to be given not just to whom to consult but when to consult them, how much, and for what purposes. Many of the stakeholders targeted by the assessment will have been interviewed or part of earlier assessments by other donors and may suffer from "assessment fatigue." In this sense, consideration regarding stakeholders should be given not just to generating political support, but also to avoid poisoning good will toward the assessment which could occur if people feel that they are being consulted, or that the assessment is being performed in a needless or duplicative manner.

For more information on stakeholder analysis, see "A User's Guide to Poverty and Social Impact Analysis" (World Bank 2003c, 49).

Choosing Assessment Methodologies

The choice of methodologies will largely be shaped by the design of the assessment, particularly where the assessment's length, scope, and cost are predetermined. Most assessments will rely on a variety of data sources (1) because no single one will provide all the information needed, and (2) because in combination, they can be used to corroborate findings and check against the biases inherent to any single method. Chapter 3 treats how to combine findings from different sources of information. Chapter 4 discusses the nature and relative merits of various soft and hard research methods, and provides pointers for using them.

For planning purposes, the choice of methods will give shape to the assessment and help consolidate the planning to that point. The preliminary identification of problems to investigate and stakeholders will help to determine which tools to use and in which context. At the same time, the mechanics of using a particular method, and its relative strength and practicality in harnessing information, will help shape the order and planning of the research, that is, what tasks can be done when. For example, to garner input from legal aid clients, a focus group discussion could be more effective in terms of cost, time, and breadth of information collected than would be individual interviews with each. Similarly, individual interviews with court administrators will help to add context to, but can also be tested or verified with other quantitative administrative data where these are available.

Box 4: Patience sometimes pays off
In a country in South Asia, the Chief Justice of the Supreme Court had many roles. Not only did he preside over his court, he also chaired the Judicial Services Commission which is in charge of managing the lower courts, and he was regarded as the head of the judiciary as a whole. Thus, there were many good reasons to interview him for the assessment of this country's justice institutions.

Initially, however, the Registrar's Office let it be known that it was not considered proper for a World Bank assessor to actually meet with the chief justice himself. After visiting most of the courts and other institutions on my list, including the registrar in question, and simply exercising patience, two weeks later, I was invited to come and meet the Chief Justice. By that time, I was able to share some preliminary conclusions at what was primarily a ceremonial visit.

In all cases, furthermore, use of the individual method will require specific preparations. For instance, individual interviews or focus group discussions require scheduling and securing the presence of the participants. The discussions also require preparing and testing, to the extent possible, a standardized set of interview questions. Similarly, trying to access administrative data requires an initial idea of how easily accessible the data are, both in terms of physical location (for example, centralized electronic files versus handwritten ledgers and haphazard filing systems) and securing the relevant permission to review the data.

Reporting and Dissemination
Writing and disseminating the assessment report mark the culmination of the assessment process. While the actual writing of the report might not begin until after the field research has been completed, constructing the report or thinking about how it will be written often begins during the research phase, as findings are evaluated and tested. The main task in writing the report is to assemble all of the findings and analysis and distill them into written form. The goal is to narrow down the data collected—choosing what to leave in and out—to present the most relevant findings in the most effective way. This exercise is part organizational, part analytical, and part political. Chapter 6 provides advice on all of these fronts. It explains both how best to organize the report's content and how to navigate political pitfalls that arise at the reporting and dissemination stage.

One factor to be considered when producing the report is that *conclusions are time sensitive*. The assessment already may be influencing the existing state of affairs before the final report is produced. The longer it takes for the report to be disseminated, the more likely the report will lose its relevance. This caveat underscores the importance, from the beginning, of planning how to draft and then disseminate the report. For planning purposes, a main task is to devise an effective system to compile information during data collection which will enable easy retrieval and analysis at the report-writing stage.

Dissemination serves a number of purposes: fostering dialogue in the country that informs reform choices, generating ownership of eventual reforms, and public accounting of the assessment. Conclusions and recommendations in the report may easily be taken as criticism. A degree of skepticism in the face of what is perceived as criticism from "outsiders" is to be expected. Involving all stakeholders in all stages of the assessment process will promote positive reception of the assessment. Suggestions on involving stakeholders are included in each chapter. Other ways to increase receptivity include the language used to present the results, solicitation of comments on multiple drafts, and the sequencing and media used to distribute the final report. Chapter 6 describes possible dissemination strategies and techniques.

3. ASSESSMENT PROCESS

The following discussion draws on the authors' experience and on a review of the body of assessments done by the World Bank, as well as examples from other donors, most notably, DFID, IADB, USAID, and the UN agencies.[12] The desk review provided excellent examples of innovative methodologies and insights. It also suggested common shortcomings for which the authors have posed solutions. These weaknesses include:

- Tendency of many assessments to stick to a largely descriptive, formalistic narrative about sector organization. While such a narrative is useful to those not familiar with the system, it probably is the primary cause of the common complaint by counterparts that "the study doesn't tell us anything we do not already know."

- Tendency to focus on the state sector, especially the courts, with minimal attention to how their functions are complemented, or even replaced, by informal, non-state institutions. Given the difficulties of reviewing or assessing the latter, this focus is understandable. However, it provides a very incomplete picture of how societies resolve conflicts and enforce rules and may yield unnecessary or counterproductive recommendations.

- Tendency to go soft on criticisms even of fairly blatant, widely voiced problems, such as corruption, and political interference. The likely negative reaction from highly placed counterparts may explain this soft-pedaling, but it puts into question the validity of the entire assessment exercise.

- Failure to prioritize the problems that are discussed so that inadequate equipment or rundown buildings get the same attention as corruption or delay.

- Inadequate explanations of methodologies used to collect assessment data, of the relative importance of different data sources in reaching conclusions, and of he standards against which findings are evaluated.

- Over-reliance on conclusions drawn from informant interviews and surveys, with little attempt to validate them with other types of data. "Hard data" (chapter 4) may be difficult to collect; but intuition, anecdotes, and conventional wisdom are notoriously inaccurate so need to be checked.

- Minimal analysis or interpretation of quantitative data, either contextually or against international standards.

- Tendency to offer cafeteria lists of recommended actions, some of which seem to have little basis in the analysis, with minimal attention to prioritization, sequencing, or urgency.

---

[12] Many of these other donors have relied on outside contractors, including NGOs such as ABA/CEELI or the National Center for State Courts; universities, such as Florida International University; or for-profit firms, such as Chechi or Chemonics.

Some of these shortcomings do not have easy remedies, especially those arising from the anticipated sensitivity of counterparts or the limited resources available for the assessment. Others might be addressed by a stronger focus on the assessment's purpose––a principal input to an eventual reform—and by closer collaboration among team members. The following discussion recommends a several-stage approach; it emphasizes problem-identification, analysis, and remedies. These emphases do not markedly change the usual assessment activities but do reorient their target. Only the final stages of comprehensive analysis and recommendations transcend the usual format because these are the areas in which most assessments seem weakest. Stage 4, an analysis of actual operations, is a partial novelty because most assessments do not consider it separately, and many do not include it at all.

Preliminary Considerations of the Assessment's Purpose and Approach

A reform program or project's purpose is to improve sector performance; the purpose of the assessment is to provide a better understanding of what is wrong and why. Determining what the problem is can be as difficult as identifying its causes or devising remedies. What people believe to be wrong may not be entirely validated by the assessment, and the findings also may indicate additional performance shortcomings.

Typically, justice reforms are requested because of perceived problems in one or more of the following areas:

- Court or other sector agencies' failure to resolve disputes fairly, effectively, and conclusively[13]
- Slow, costly processing of "cases"[14]
- Limited access to sector dispute resolution and related services, especially for the poor, women, or other marginalized groups
- Corruption or politicization of services
- Biased decision-making
- Impacts of one or more of the above on such extra-sector goals as economic growth, poverty reduction, and citizen security and civil peace.

The positive values implicit in these complaint statements provide a preliminary definition of the characteristics of a well-functioning system: timeliness, honesty, broad and equitable access, and impartiality. These same values also are reflected in an increasing number of international conventions as well as in statements by the World Bank and other donors as to their objectives in supporting justice sector reforms.

---

[13] "Other agencies" may include mediation centers, administrative tribunals located outside the judiciary, and various traditional dispute resolution bodies. They also include certain auxiliary agencies (police, prosecutors, registries) that provide input to the judicial process or are authorized to make specific decisions (in the case of police or prosecutors in some countries, to levy fines or impose community service), thus diverting disputes from courts. Whether their performance in either area is covered in the assessment depends on the problems being investigated.

[14] "Cases" here denotes the basic output of each agency, which may be a judgment (for the courts), an investigation (for the police), or a title (for a registry).

However, it may be still more significant that these are the problems mentioned most often by national citizens. This fact suggests that even if a full-blown sector model is missing, citizens hold certain widely shared expectations as to what the sector does and how it does it.

These problems are listed above in their broadest forms, and it is not uncommon for an assessment to investigate several or all of them. Alternatively, the assessment may be directed at a subcategory of only one problem. Examples are delayed contract enforcement in debt collection cases heard by civil courts, the judiciary's impact on credit availability, ineffectual handling of cases of alleged administrative corruption, or the impact of judicial services on poor women seeking child support.

However broadly or narrowly its initial aims were defined, the diagnostic's approach should be institutional. In other words, it should focus on the formal and informal norms, organizations, actors, and their incentives as they shape justice sector outputs.[15] Many assessments use a detailed analysis of the legal framework as their point of departure in addressing these questions. However, starting with the existing legal system generally is not a good idea. Existing legislation is important in understanding official aspirations and some constraints on system performance, but it is not everything. If it were, the considerable efforts invested in changing laws throughout the developing world would have produced far more positive results. In fact, the law is one of several institutions that shape sector output, but to understand why output is perceived as inadequate, much more must be considered.

This handbook defines the justice sector or system most broadly as:

> *The institutions that are central to resolving conflicts arising over alleged violations or different interpretations of the rules that societies create to govern members' behavior; and that, as a consequence, are central to strengthening the normative framework (laws and rules) that shapes public and private actions.*

This definition is the one most commonly used by the World Bank. Admittedly, it is a mouthful. However, it is necessarily so to encompass the variety of institutions involved in rule-based conflict resolution[16] and to enable the inclusion of the non-state organizations and rules that may serve these purposes for a greater or lesser portion of a nation's population.

Justice, or legal and justice reform, concerns not only courts and state law. Nevertheless, the assessment ultimately may emphasize that part of the universe because it is most relevant to the problems addressed. The assessment also may focus elsewhere, especially

---

[15] This should not be conflated with the neo-institutional or law and economics approach, which while also "institutional" in its analysis, incorporates additional assumptions about the sector's impacts on economic transactions. Some assessments (albeit none reviewed here) adopt this theoretical framework, but an institutional analysis does not require this step.

[16] Examples on the formal side are the courts as well as the legal profession and subcategories including notaries, prosecutors, and public defenders; registries; arbitration centers; and police.

in countries in which the formal state institutions have an extremely limited reach and the conflicts they do not cover contribute substantially to social unrest or negatively affect the poor. Thus, at least in the initial stages, it is best to keep options open, narrowing them only once additional decisions are made.

Getting Started: Defining the Focus and Scope

Of course, the disadvantage to an institutional approach is that it does not establish an equally clear point of entry as starting with existing law. If the assessment does not start with the law, where does it start? There are two answers, both dictated by circumstances. The first is to focus on understanding a particular predetermined problem. This problem may be system specific, such as delay or insufficient access. Alternatively, it may be extra-system, involving the sector's presumed negative impact on developmental aims—poverty reduction, economic investment, or criminal and political violence.[17]

A second option would be to start the assessment with a more open-ended search for problems and their causes. The reason is that the "problems" first identified turn out often to be less important than others, or sometimes nonexistent.[18] Either starting point lends itself to an institutional analysis. The only difference is that between the search for answers to a predefined question, or the search for questions and then the answers.

For Bank staff, the choice is likely to be made by the Country Management Unit (CMU) or the country assistance strategy (CAS). Based on the CAS discussions, the CMU will request a project focusing on themes such as "judiciary's impact on contract enforcement," or "poverty reduction," or a "justice reform project," with no additional instructions. If the assessment team has the flexibility of this second option, it is advised to begin by canvassing opinions to define the major problems. The results will not be the final problem statement, but they can help in reaching it. As with the rest of the assessment process, determining an assessment's focus is inherently iterative, and problem definition may be the most iterative aspect.

To avoid unnecessary confusion, the essential distinctions among *problems, causes, effects,* and *remedies* or *solutions* should be kept in mind. There are various ways to highlight these distinctions, but in this handbook, these terms will be defined as follows[19]:

---

[17] Although criminal justice is presumed to affect crime levels, the many other contributing factors suggest that the incidence of crime should be considered a downstream or extra-system value on which the justice sector has an impact.

[18] This was the experience of ESW contracted by the Mexico CMU to investigate the impact of the judiciary on contract enforcement, especially in terms of excessive delays and pro-debtor biases. As it turned out, the creditor won more than 90% of the time; the delays were not in the prejudgment phase but in enforcement, and they were largely attributable to factors outside the judiciary's control. Moreover, while the initiative started as a request from commercial bankers (interested in proving that the courts were responsible for high interest rates), banks were found to be minority users of the courts and to almost always win their cases. Any negative impacts of slow or inconclusive adjudication fell more heavily on individual users and small businesses. As for interest rates, anyone familiar with Mexico knows the courts are a minor contributor at most (World Bank 2002).

[19] The terminology used here is arbitrary, but the conceptual distinctions are vital.

- *Problems* will be reserved for unsatisfactory system outputs, such as delays, corruption, or limited access.
- *Effects* or *impacts* are the downstream societal consequences of these problems on extra-system values such as economic growth, poverty reduction, and political stability. The resolution of problems or improvement of effects provides the motive or justification for a reform and thus for the objective to which the assessment contributes.
- *Causes* are the reasons for the deficiencies named "problems" and "effects." Causes are of two types. A *proximate* cause is the immediate factor that seems to trigger the problematic performance, for example, low judicial budgets and salaries, under-qualified judges and staff, or poorly drafted laws. The *underlying* cause is the more basic, structural explanation for both the problem and its proximate cause. Examples are resource poverty for the country as a whole, politicians' desire to control the courts, or an experience-based expectation that laws will not be enforced and thus that their content is at best aspirational.
- *Remedies* improve outputs and impacts by attacking proximate and/or underlying causes.

The reason for stressing these distinctions is that many reform strategies seem to confuse or conflate the categories. It is not unusual, for example, for judges and ministries of justice to identify *candidate* proximate causes (not enough computers or training) as problems and to propose the provision of the missing element (the computers, the equipment, the infrastructure) as the solution.[20] This apparent tautology is not entirely illogical. From the respondents' standpoint, insufficient equipment is problematic in that it can make their work more difficult. However, in terms of the above categories, it is only a potential proximate cause of the problems that really interest the analysts and system users: the quantity or quality of output.

Counterparts can be quite adamant about their analyses of a situation. Therefore, it is very important, even in the early stages of an assessment, for the team to (1) ensure that interviews about problems reach beyond high-ranking judicial staff and political leaders and (2) validate their initial appreciations, when possible, against empirical data.[21] Validation also guards against basing strategies on unrealistic expectations or erroneous perceptions. For example, system users may believe that every case should be resolved in a matter of weeks or simply may be misinformed as to the real times that most cases take. Conventional wisdom can be very powerful, but it often is in error. A good assessment is similar to a good medical examination. The latter often starts with the patient's

---

[20] "Candidate" is used because, for example, it is not always clear that too few computers are a proximate cause of a problem (such as delay).

[21] If validation is not possible, the assessment and final report should recognize them as hypothesized weaknesses, noting as well the extent to which the view is held, the types of groups or actors holding it, and the additional measure that might be used to test its validity. Assessments are likely to turn up many contentions of this type, and even if ultimately untrue, they are important in shaping how individuals deal with the legal system. Even if the problem (for example, judicial corruption) is far less prevalent than believed, the belief in its existence can be an explanation (cause) for other problems (for example, limited access or noncompliance with judgments).

complaints and self-diagnosis, but it may reach quite different conclusions as to what is wrong and why.

Whether it begins with a predefined problem or works off a short-list gleaned through preliminary interviews, the rest of the assessment has basically the same format. What is emphasized in each area depends on what is already known and what the team is seeking to do: whether it is, for example, exploring a list of candidate problems or investigating a predetermined topic to identify causes and alternative solutions

The discussion below and the outline in appendix 1 focus on the assessment process, rather than on the assessment as final report. This is another important distinction. The assessment team doubtless will collect more information than it will include in its final written product. In fact, trying to include all of the information is another common error. Part of the information collected is simply what team members need to better orient themselves. Part is essential to test the working hypotheses as to problems and to provide support for the conclusions reached. After the final analysis is done and recommendations are developed (stages 5 and 6 below), the team must make decisions as to what will be left in the report, what will appear as appendixes, and what will remain in the background papers or be discarded (chapter 6).

Assessment Stages

The following discussion focuses on the stages in the assessment process, which often coincide with the chapters in the final assessment. While the intent here is not to provide an outline for the written report, suggestions appear below as to how information might be incorporated in it. The increasing number of assessment guidelines developed by other agencies often supply checklists of questions to structure fieldwork. While reviewing several of these lists may help teams organize their work, this handbook does not attempt to follow the checklist model because:

   a. Checklists inevitably include an impossible number of details, many of which will be irrelevant for a more focused exercise.
   b. The cafeteria approach provides little guidance as to how to process what is collected, either the significance of individual responses or their collective use to reach prioritized conclusions.
   c. In compiling checklists, it is extremely difficult to avoid inserting the authors' disciplinary and ideological biases.

Most checklists tend to stress formal rules and basic, largely quantitative data on resources, for example, the number of employees, cases, and computers, or the size and condition of infrastructure. While much of this data may be collected in stages 1 and 3 below, the format outlined here places more emphasis on problem definition (stage 2) and its role in orienting data collection, on the analysis of real outputs (stage 4), and on two additional steps: the formulation of composite conclusions and of prioritized alternative recommendations.

Stage 1. Preliminary background

All assessments start with a "land, people, and justice sector" section that lays out basic details critical to interpret the rest of the material and also to decide on recommended courses of action. While this information can be gathered at any stage in the process, it will be useful for the team to collect and analyze some of it first. In and of itself, this information can provide initial ideas of where problems may lie. It also will be extremely helpful in interpreting the rest of the data collected and in identifying key stakeholders, informants, and partners who might collaborate in the research. Much of the background material can be taken from existing documents and usually can be reviewed before the team ever gets to the field (chapter 2).

This background should include at least two parts: (1) the political socioeconomic situation, and (2) the basics of the justice system. A suggested outline of key points to be considered is provided in appendix I, along with those for the rest of the assessment. Preliminary material on the justice system should include a brief overview that covers both state and non-state normative (laws and rules) and organizational frameworks, principal interactions among the components, recent changes, and major reforms under consideration. This material is useful even for assessments with a predetermined narrow focus. All or the relevant parts of this material can be fleshed out later with the field work, but having an initial sector map can be an enormous help in deciding where to focus subsequent research. Either in this or the country background portion, the team should attempt an initial listing of the most common and most important types of conflicts characterizing the society and the extent to which the justice system addresses them.

Stage 2. Preliminary problem identification

Stage 2 begins the real assessment and the collection and analysis of information that constitute its value added. Aside from whatever instructions are provided by those requesting the assessment (WB, CMU), preliminary problem identification has two sources:

1. What people say, including information from written documents, existing surveys and opinion polls, and international ranking systems[22]
2. What the assessment team observes based in part on comparisons with the performance of other systems.

It is with the second source that expertise concerning justice systems elsewhere begins to pay off. A substantive expert with broad comparative knowledge can quickly identify anomalies that would be invisible to local experts who lack that comparative base. Many judiciaries that handle remarkably few cases claim overwork, and this impression may be

---

[22] These ranking systems include composite scales such as the WBI rule of law (and other governance) indicators and independent indices such as that produced by the Bank's Doing Business database (WB 2006). While both databases have their critics, they are nonetheless very powerful tools for identifying potential problems.

shared by citizens.[23] The same is true of statements about excessive or insufficient budgets, delays, and numbers of judges, courts, and support staff. There are no hard and fast rules on these statistics, but there is a normal range of variation for each that can be used to guide initial hypothesis formulation.[24] The increasing availability of international databases provides one set of standards against which to assess reported or observed problems. These databases also may be useful in fleshing out the more qualitative norms expressed in international conventions. Their application need not be limited to the obvious values such as delay or reasonable workloads. With a little creativity, they also could be used to explore qualitative issues including independence, access, or impartiality. The previous statement is not an invitation to convert the assessment process into an exercise in quantifying everything. However, when dealing with highly sensitive topics, some quantitative comparisons, such as turnover rates on the bench as an indicator of independence, can bolster what otherwise might be received as ideological or ethnocentric claims.

The results of these problem-identification exercises should be regarded as working hypotheses to be tested and refined against the findings from additional fieldwork. Teams may find it useful to rank problems and to associate them with different groups. The ranking and even the problems identified may vary considerably depending on the source. This stage also can include informants' and team members' preliminary identification of causes, which may be useful for the later work. Even if not validated by the rest of the process, these initial appreciations (or at least those from local sources) should be included in the final report to show that they have been taken into account and to allow a presentation of the counter-evidence.[25]

Stage 3. Collection of information on targeted sector organizations and actors: Resources, capabilities, formal and informal rules, and incentives

Stage 3 is the meat of the assessment and incorporates the majority of the fieldwork. How much information is collected, how its collection is organized, and how much emphasis is placed on each part depend (1) on the problems targeted and (2) on the human, financial, and temporal resources available. This handbook assumes that contemporary assessments will not have the luxury of the $250,000 budgets, 40-person national and international teams, and a year or more of time that USAID invested in the assessments done in Central America in the late 1980s (when, it bears mentioning, a quarter of a million

---

[23] Several pieces of ESW conducted in Latin America have explored this topic in passing. See World Bank 2002; 2003a and b; 2005). On closer examination, even some apparently overloaded Latin American judiciaries, such as Brazil's, have been found to have a large number of fairly simple cases (summary debt collection or the correction of common administrative errors in readjustments to pensions) that did not require much attention from their judges.

[24] For example, if judges average 8,000 filings per year (far above the normal range), it could be surmised that an excessive workload could lead to problems such as delays. However, if the average filings are under 200 (below the normal range), delay (which may occur, but must be established separately) is unlikely to be a result of the sheer volume of workload.

[25] For example, in ESW on debt collection in Mexico, these initial complaints (most of which were invalidated) were used as working hypotheses and were discussed in the final report, which contrasted the conventional wisdom about judicial performance flaws with what the data revealed. (World Bank 2002).

dollars went a lot farther than it does today). These generous terms, which also characterized some early WB assessments, allowed for extremely comprehensive studies.

It is more likely that present-day teams will have a few weeks for fieldwork, a few more for the preparatory and final stages, and a budget far lower, even in nominal terms, than that mentioned above. These conditions argue for much greater selectivity of focus from the start. They also mean that both the team and those contracting the assessment should be clear as to what the resources will allow.[26]

Sometimes, the scope of the assessment is extremely narrow; the team might be asked to focus, for example, on delays in processing debt collection cases or obstacles to using courts affecting poor citizens in X country. Except for such situations, it is probably most practical to structure fieldwork around a one-by-one analysis of the major organizational actors and to deal with their staff or members within this context. In fact, even for assessments with a predefined problem focus, this structure still may be useful but in an abbreviated form.

As detailed in the outline (appendix 1)**,** state or formal sector organizational actors commonly include the following:

- Ordinary judiciary; when relevant, separate constitutional or administrative tribunals; and where recognized, religious or other special courts
- When criminal justice will be considered, other criminal justice actors: prosecution, defense
- Private bar
- State attorneys
- Notaries
- Property registries
- Formally recognized arbitration services
- Other legal aid providers.

For these agencies and groups, the following categories should be covered in greater or lesser detail:

- Details of organization (again a chart may be most useful) and of overall powers and duties, as well as internal distribution of labor
- Body(ies) responsible for organizational governance and administration: composition, powers, focus of operations (day-to-day administration, policy setting, planning)
- Human resources: Overall number; major job categories and distribution of work force among them; employment conditions (salaries, tenure, career system, or not) means and conditions for selection; performance monitoring, disciplinary

---

[26] The authors have avoided suggesting monetary amounts because so much depends on the quantity of information already available, size and heterogeneity of the country, scope of the work, methodologies used, nature of the problems explored, and local cost structures.

procedures and removal or disqualification, if any; skill levels and training programs

- Geographic distribution of work units, employees, and workloads
- Budgets: Sources, how set, how managed (and by whom), functional and geographic distribution
- Other resources—infrastructure, ICT equipment, vehicles—and their distribution
- Laws/rules ("normative framework") governing operations (including any performance standards and required release of information), process required for changing them, and brief summary of their known or likely impact on real operations
- Rules to access their services, likely impact of the above on access (geographic distribution, payment of fees, need for legal representation).

This is a short list of suggested topics. However, any one of them could keep the designated team specialist (in legal matters, ICT, human resources, or financial and administrative management) busy for months. The key is *selectivity*: a quick reconnaissance to collect sufficient information and, where available, statistics, to test working hypotheses and support emerging conclusions regarding weaknesses, but not an exhaustive review of all data. Shortcuts also can be taken by prioritizing certain agencies, most probably, the courts, and doing a much more cursory review of the others. Some groups of actors that are included in the organizational review either lack a formal organization or do not operate under the organization's direct control. Examples include notaries, private bailiffs, and private attorneys. For them, some categories simply may be irrelevant.

For nonstate conflict resolution bodies and other informal "organizations," the review will be similar but may be partial in its coverage. This is true, especially in cases in which a multitude of different, geographically limited structures exist. Additionally, some of the above categories may be irrelevant or have to be interpreted contextually, for example those on career structures, budgets, or "other resources" In these cases, the team may also add questions, for example, on these bodies' official recognition by and linkages with state institutions. Should the assessment's conclusions point to a greater emphasis on these organizations, work with them will obviously require another, in-depth assessment of their structure and operations.[27]

World Bank and other donor assessments increasingly recognize the importance of traditional or community conflict resolution mechanisms, but do not review them in depth and offer no methodologies for doing so. The few existing standalone studies (see Cooter 1989 on customary land law in Papua New Guinea) were very labor intensive and thus are no guide for a quick reconnaissance. Possibly, the heightened interest in non-state institutions will produce techniques that assessments can incorporate.

---

[27] The authors are assuming that the initial instructions are not to focus on nonstate entities. If they are, those requesting the assessment need to consider two caveats: (1) the time and costs will be far greater owing to logistical problems and the likely lack of preexisting information, and (2) different and more costly methodologies will be required to collect the basic data.

One topic overlooked in most assessments but critical to this and the following section is the role of auxiliary organizations and informal institutions in complementing the actions of key state agencies. For example, the enforcement of judgments or the problem of contract enforcement in general hinges in part on the presence of effective property registries, credit bureaus, and bailiffs, *huissiers*, and other enforcement agents. They also depend on numerous informal rules and practices present even in the "modern" sector. In a study of contract enforcement in Brazil and Chile (Stone and others 1996), the authors argue that Brazilian entrepreneurs' reliance on reputation and informal networks brings them results comparable to those in the more court-centered Chilean system. Conversely, in summarizing Hendley's (1999) work on Russia, Dethier (1999, 40) concludes that the efficacy of the Russian court system is undermined by cultural norms (the tradition of nonpayment of debts) and "few reputational sanctions."

An overview assessment could explore, however briefly, the role of formal auxiliary actors and agencies, but it is doubtful that it would go far in identifying these informal norms. Nonetheless, team members should be aware of the potential impact of informal norms in order to avoid recommending "reforms" that these norms make less necessary or that will not work unless exogenous practices and values also are changed. .

These cautions are especially relevant for the next step, the overview of sector operations, and may be more easily addressed there. Trying to identify cultural values in the abstract may be as futile as trying to analyze the "goodness" of law out of context. However, when the team turns to explain specific outcomes, the narrower focus makes more feasible both types of analysis.

Stage 4. Summary overview and analysis of operations of sector organizations

A certain number of tentative conclusions can be derived from the review of the composition, legal mandates, real powers, and resources of sector organizations/
operations. However, it is desirable to check these conclusions against what happens on the ground. This step is most often omitted in standard assessments. The best means for doing this analysis of random samples of "cases" is too expensive and time consuming for the types of assessments described here.[28] In some countries, academics or even the organizations themselves already may have done such work, in which case it can be incorporated in the assessment. Unfortunately, that is a fairly rare situation.[29] If organizations have kept any kind of statistics on their operations, these data may be used to develop an aggregate snapshot of performance: quantities of "cases" handled, rates of clearance, outcomes, and trends and differences in all by geographic or functional area.[30]

---

[28] The authors retain the quotation marks because, for some organizations, "cases" may not be the correct term. What is meant is the basic unit of output, whether an investigation, a client defended or provided with information, or a title registered.

[29] In Latin America, for example, aside from WB-sponsored research and some spin-offs that it has inspired, the only countries in which such studies are available are Brazil and Colombia. Moreover, these studies are either outdated or very limited in their coverage.

[30] For an idea of what can be done even with fairly rudimentary national statistics of a somewhat dubious quality, see World Bank 2005. Figures 1 and 2 in chapter 4 give two examples. It should be noted,

This aggregate information can be supplemented by observation of case processing or examination of a few records. Thus, a recent Inter-American Development Bank (IADB) study in Mexico reviewed records of 80 judgments reached in 3 state court systems.[31] The sample was not scientifically drawn and looked at only a few trial courts in each state capital, but it was done quickly by 2 team members and provided valuable insights. Such supplementary information is to be regarded as illustrative, not representative. Nevertheless, it can be used to check for bottlenecks, for biases in results or in what gets into or through the system, and for problems originating in the required inputs from a number of organizations or organizational actors.

For example, in criminal cases, bottlenecks frequently originate in the police or prosecutorial investigation, or in inadequate coordination between the two. Delays may be party generated, involving a proactive defense or either party's failure to meet deadlines or court appearances. In noncriminal cases, delays also may result from the failure of the plaintiff to request forward movement. Certain procedural stages— notification or assets seizure—typically pose problems, either because of logistical obstacles or inadequate supervision of the responsible officials. One frequently overlooked bottleneck in analyses of civil and commercial cases in particular is enforcement of judgments. Tracking of processing should try to capture this as well.

The kinds of cases selected for this review will be dictated by the problem focus—rights, enforcement, child support, debt collection, bankruptcy, or some other topic. Aggregate or selective review of disputes resolved by traditional and other non-state mechanisms will be more difficult, because they are less likely to have statistics or written files and may not welcome outsiders' examination of their decisions. The team may substitute interviews with users may, although even here a problem of access or simple logistics is likely to interfere. The assessment may have to work with a few illustrative examples or depend entirely on what studies or written documents already exist.

Once information is collected, three kinds of further analysis can be attempted:

1. Graphic representations of case trajectories, noting different outcomes and likely bottlenecks, barriers, and biases.
2. Preliminary effort to relate performance variables to the normative framework (rules and laws) and resource endowments. Much of this material is collected in stage 3.
3. Identification of key actors whose inputs further shape the outcomes in ways not entirely attributable to the formal norms; and of the factors (skills, other resources, external controls, and probable incentives) accounting for these differences. (See chapter 4 for a brief discussion of incentive analysis.)

---

however, that the process was facilitated by internet access to national and subnational databases. Even then, it took several person-months to collect, interpret, and process the information.

[31] The study is not included in the References because it has not been released.

It also may be useful at this stage to develop a tentative map or schematic chart of conflict resolution, showing at which point issues enter and are resolved, who accesses which entities, and which types of issues and clients seem to be excluded from treatment by any dispute resolution mechanism. At this point, the results from step 1 (outlining the major conflicts characterizing a society) and the interviews in step 2 (defining problems) can be referenced. While the most tentative of all (especially regarding what is not covered), the mapping exercise nonetheless is critical to address both access issues and questions about impacts on levels of social conflict, poverty reduction, and economic growth

Data collected here and their initial analyses are intended to link the information collected in step 3 on system characteristics to actual outputs, and so to problems already identified in step 2, or revealed and/or elaborated in step 4. The output analysis thus is critical to comprehend the real problems and identify their causes. Generally, the resources available will not permit a fully representative picture, nor even an investigation of all problem areas. Rather, the data and analysis from step 4 will provide the best means of testing initial impressions and conventional wisdom about performance failings and their causes.

Stage 5. Comprehensive analysis of partial findings, prioritization of problems and their causes, and identification of areas for interventions

The findings and partial analyses comprising the first four sections normally will generate:

- A list of performance and impact problems
- A variety of interrelated causes
- Initial ideas for solutions.

The challenge lies in consolidating and prioritizing these initial sets of conclusions. Even so simple an assignment as explaining the delays in certain types of cases or identifying the barriers to the poor faces this challenge.

Either starting point doubtless will identify a series of proximate and underlying causes and an initial idea as to which are:

- Most important because of the significance of the problems to which they contribute, for example, high incidence of perceived or real corruption, escalating levels of crime or social conflict
- Most urgent because the problem demands immediate resolution, for example, a vital service that is about to disappear due to lack of funding or other attention[32]

---

[32] One example from Peru was the near abandonment of the state-of-the art judicial archives created under the Fujimori administration (1990–2000). With this kind of service, six months of inattention would have undone 5 years of work and a several-million-dollar investment

- Susceptible to easier resolution, thus generating quick wins, for example, providing services to populations in urban slums by buying a few buses to create itinerant courts.

However, where the assignment was more open-ended, the challenge can be enormous. Perhaps there are problems of delays, access, unfair decisions, high costs (for users and for the state), and corruption. *Proximate causes* may include deficiencies in the law; poor quality, motivation, and distribution of staff; absence of staff development programs; low and poorly utilized budgets; failure to monitor performance; poor record-keeping; various procedural bottlenecks owing to inadequate inputs from actors within the same or different organizations; and a variety of physical, financial, and cultural barriers to marginalized groups.

Underlying causes may include limited managerial capacity within the sector (or the country), high levels of political interference in internal operations, various legal restrictions on human and other resource use, deficiencies in the overall quality of the pool from which staff is drawn, governments' own financial limitations, population dispersion patterns, ethnic divisions and conflicts, levels of poverty and inequality, and poor law drafting capabilities.

Existing assessments offer little specific guidance on how to consolidate this data, and many do not make the attempt. The following suggestions and those in the next section are the authors' best approximation of what might be done in a problem-focused process. Although reforms usually work on the proximate causes, some may be less susceptible to change than others because of the reasons underlying their existence. If budgets are low because the country is poor, this underlying cause will not be changed rapidly.

Moreover, any one problem is likely to be backed by *several* proximate and underlying causes. The truism that delay or inefficiency derives from poor incentives and overly complex procedures hides a multitude of contributing factors at both causal levels.[33] The team may be tempted to start analysis with the underlying causes, which ultimately may be fewer than the proximate ones. However, starting from the opposite end with the problems is usually more practical. Doing so certainly will be the most diplomatic way to present the analysis. Following from this suggestion, five additional notes on analysis can be added:

1. Consolidate the candidate list into a few principal problems and rank them by importance by using stakeholder surveys from step 2, additional instructions from the agency requesting the assessment, and implied downstream impacts.
2. For each of these problems, list and rank the proximate causes and link them to the underlying ones.
3. Cross-check the lists to identify areas of overlap at the proximate and underlying causal levels, using the exercise to develop a schematic representation of problems and interrelated causes.

---

[33] Botero and others 2003.

4. Based on the above, first distinguish (a) problems lending themselves to targeted improvements (because of the greater ease of addressing proximate causes, the lesser impact of difficult underlying ones) from (b) those requiring more complex, longer term remedies. Then identify proximate and underlying causes that, because of their impact on several problems areas, merit greater, if longer term, attention.

5. Finally, use this last step to outline a series of potential points of entry for change, noting their likely short- and longer term results, the difficulty of achieving them, and the sources of support and opposition for implementing them.

This comprehensive analytic overview completes the data collection and analysis. The results will be a first approximation of the problem or problems posed by the systems' operations and the identification of their interrelated causes.

Stage 6. Recommendations for reform programs

The above exercise, in which the entire team will participate, is the basis for a reform strategy. The assessment focuses on problem resolution and is embedded in an analysis of the principal areas of poor performance, their causes, and the means recommended to address them. Ideally, and contrary to the usual practice, recommendations should have a short-, medium-, and long-term focus. While including some quick fixes, or at least targeted improvements achievable over the short run, the assessment also will incorporate the identification of areas requiring concerted attention but not amenable to rapid change.

Quick fixes might include changes to a law or procedure to eliminate obvious sources of delay. Examples are the substitution of an oral hearing for a lengthy exchange of written documents, or of a claims-filing form to facilitate processing and eliminate time-consuming irrelevant additions; or the reorganization of notification services. Other forms of quick fixes might enhance access. They could include the elimination of filing fees for certain types of cases or clients or the introduction of court interpreters or of mobile courtrooms. Staff training and public education programs also can be useful so long as they are targeted at resolving specific problems, such as informing citizens how to access specific services or correcting errors in the application of laws or procedural requirements. Most quick fixes still require a year or two for adoption. Nevertheless, where introduced, even on a limited geographic basis, they can produce visible and measurable improvements in output and also demonstrate the feasibility of change.

However, quick fixes should be seen as a reform tactic, not the ultimate solution for basic problems or their causes.[34] Basic problems generally will be addressed by medium- and long-term recommendations, based not only on findings about performance and impact

---

[34] The short-, medium-, and long-term focus hinges on the ease with which contributing causes can be attacked. However, their application also requires a prior ranking of the importance of the problems addressed. Except to demonstrate the possibility for change, there is limited value in quick fixes for unimportant problems. In any event, if delay is a principal problem, the quick fixes mentioned here are unlikely to eliminate it, but they should make some measurable difference over the short run. It also bears mentioning that absent the medium- and long-range strategy, these initial improvements may well not be sustainable.

weaknesses but also on an evaluation of the various contextual obstacles. The medium- and longer term actions likely will be linked. Medium-term strategies can lay the groundwork for addressing problems that cannot be fully resolved even within that timeframe, and they should not in any case propose measures that would impede longer term solutions.

For example, one proximate cause for many performance problems may be the poor quality and excessive or insufficient number of support staff. A medium-term strategy might focus on improving support staff already in the system (via training, monitoring, performance evaluations), while a longer term goal might be to create a separate career for these individuals. Depending on local circumstances, the creation of a career could be a medium-term goal. However, in many countries, doing so would be impossible because of such factors as secure tenure, which prevents incumbents from being removed or even reassigned, quality of the pool of recruits, weight of political patronage, or simple budgetary constraints. In a similar fashion, if the long-term objective is a state system that reaches all citizens, the medium-term goal may have to rely on improving alternative services while considering how they might eventually be better linked to the state system.

This section will not lay out full-blown strategies but should identify likely short-, medium-, and long-term objectives with a series of alternatives to meet them. The team should draw heavily on experience elsewhere, especially regarding implementation obstacles and timelines. It bears emphasizing that short, medium, and long terms are not (as often appears to be the case) immediate, 1-year, and 3-year goals. Instead, this handbook describes goals that could be accomplished in 1 or 2 years versus those that might begin now but will require 5 to 10 or more years for reasonable, incremental implementation.

4. DATA COLLECTION METHODOLOGIES

Most assessments will rely on a variety of data sources because (1) no single one will provide all the information needed and (2) in combination, they can be used to corroborate findings and check against the biases inherent in any single method. Two important rules, almost universally honored in the breach, are to (1) incorporate in the written report a summary description of the methodologies used and (2) link findings and conclusions to the data sources on which they are based. Most assessments do provide appendixes with a list of individuals interviewed and documents consulted. Most also footnote sources of quantitative data.

However, a more explicit approach would enhance credibility. It would note, for example, that findings on court operations were drawn from interviews with 30 judges and 15 lawyers and observation of proceedings in 3 district courts. It also would indicate that conclusions about excessive or insufficient workloads were based on the team's analysis of aggregate data provided by the central statistical office and evaluated against comparable international statistics published by the European Commission (CEPEJ) or the Latin American Center for Justice Studies (CEJA).

In contrast, the usual implicit approach either presents data without interpretation or assumes that readers will realize that most findings draw on interviews, tends to invite criticisms and inferences contrary to the team's own conclusions. Respectively, the responses might be, "How do you know that?" and "My, the courts are clearly under-funded if they get only 6 percent of the national budget." An assessment is not expected to stick to the strictest standards of empirical research or to footnote every statement, but it should explain how its conclusions were reached. The most practical means for doing this is to provide a brief summary of data collection techniques in an early section and judiciously add footnotes or explanations in the body of the report. More sophisticated quantitative methodologies will require their own appendixes, discussing, for example, how surveys were conducted.

Conventionally, data sources can be divided into soft (more qualitative or subjective) and hard (quantitative, objective) categories. The traditional "soft" sources are:

- Documents, which, in addition to published studies, include laws, unpublished reports, and, more recently, material available on the internet publications
- Open-ended or informant interviews
- Observation.

Among the "hard," or more quantitative, sources are:

- Statistics already collected by the relevant organizations
- Statistics or data on organizational outputs or procedures collected by the research team via sampling of case files and similar methods
- Data from existing surveys

- Data from surveys or multiple, structured interviews conducted specifically for the assessment.

In addition, as discussed below, there also are newer, soft sources for data collection and analysis. These include the use of focus groups and similar multiparticipant exercises, evaluation of judicial decisions, and incentive analysis.

Despite an increasing emphasis on more quantitative data, soft sources continue to provide the major inputs to any assessment. Because they are less structured, soft sources usually generate more varied kinds of information and insights. The problem can be to establish their validity and reliability, especially when they produce unwelcome findings or when they are reviewed by readers more accustomed to working with statistical databases. In addition to convincing the doubters, quantitative or "hard" data are important as a means to test some propositions emerging from the softer sources. There is no better way to determine whether "everyone" really distrusts judges or whether delays are "lengthy" than actually doing the measurements.

However, hard data are expensive to generate, and what already exists may not be organized in a fashion enabling investigation of key questions. For example, courts frequently keep statistics on annual cases filings and dispositions. However, they often do not disaggregate them by the type of case and usually do not include information on delays or on the form of disposition (judgment, dismissal, or withdrawal of the complaint). Thus, as elaborated below, to explore these issues, the team will have to supplement existing statistics with other, usually expensive, methodologies.

In addition, both the collection and interpretation of hard data benefit from the kinds of insights only soft sources can provide. In conducting a survey, it is always better to collect considerable background from interviews and documents on system operations and what informants commonly identify as problems before designing the survey itself. Standardized questionnaires are useful for cross-national comparisons, but they are far less helpful in understanding what is occurring in the system under study. Hard data also require careful treatment to ensure validity—that they measure what they purport to measure—and reliability—that they are always measuring the same thing. Even when statistics are available from the courts or other sector agencies, it is important to determine how these data are generated, what the categories mean, and what kinds of checks are made on data entry. Generally, in this sector, performance statistics are a relative novelty. Their collection entails many problems that the team should be aware of. The discussion below will help the team with addressing the problems they may encounter.

No data source serves all purposes. The following discussion of different data sources begins with the conventional ones (points A–D), covers recent, more quantitative additions (E–G), and ends with a discussion of a few less conventional subjective (soft) techniques for data collection and analysis. This handbook avoids ranking data sources according to what is most desirable, because what is used depends on the topics being investigated, what is available, and the limitations on time and resources. Despite

individual or disciplinary preferences, most assessments will require and benefit from a mix.

Conventional Soft Sources

A. Documentary Sources

Within the last several years, much has been written about particular legal systems and their problems. In some cases, background material on the basic contours of a judicial or legal system is readily available. For example, Herbert Kritzer's collection, *Legal Systems of the World: A Political, Social, and Cultural Encyclopedia*, is an excellent starting point. Beyond the growing literature on the performance of legal and judicial systems and the numbers of country studies from both international and domestic sources, much information is available online. Many countries maintain their own official websites on the basics of their legal systems. These sites usually can be located on the internet by referencing the country name and "legal system" or judiciary. Often courts, ministries of justice, and other sector agencies also have their own websites. There also are regional sites kept by universities, nongovernmental organizations (NGOs) and research foundations. None of these sources may be comprehensive, and they often are not up to date. Nevertheless, even at their most limited, they provide a valuable piece of information. When possible, collecting and synthesizing all prior descriptions and evaluations in a preliminary desk review can save considerable time and money during the field research stage.

However, circumstances are not always conducive for pre-mission desk (or cyber-) reviews. First, the amount of publicly available information on any given country varies greatly. For example, anyone seeking background material on Argentina or Guatemala will find more than can possibly be absorbed in the few weeks prior to the field work. At the other extreme, detailed treatments of the legal and justice systems in Lao, Mali, or even China are in short supply.

Second, many relevant documents (particularly those that go more in depth or contain empirical data) are inaccessible and will have to be obtained on site through personal contacts and persistence. It should be noted that many donor studies never are processed in final public form so must be sought in central agency or field mission archives or in the authors' personal records.

Third, pre-existing assessments may be of varied quality, hard to evaluate, dated, and thus unreliable. There is an unfortunate, but marked tendency, for later evaluators to simply reproduce the contents of earlier studies. Consequently, their information often is far older than the drafting date suggests. Therefore, the assessment team must decide how much time to invest in tracking down documentation on the systems it is studying and take these inputs as useful starting points that are well short of authoritative conclusions. If resources allow, it is advisable to hire a research assistant to track down what is readily available and provide the team with a preliminary list. Team members then can either decide what they will try to cover or divide up the responsibility for an initial quick

review, thereby identifying the sources on which they will individually or collectively spend more time.

B. Analysis of Legal Documents

Although also publicly available, laws, constitutions, and other elements of the legal framework are treated separately because they pose different challenges for the assessment. Clearly, they should be reviewed early on (as part of the desk study, if possible) and referred to later as questions arise as to their application and impact. However, detailed law review might best be reserved for a later, even post-assessment phase. It is important to know what the law establishes as the formal practices, but this will not be enough to understand reality on the ground. Again, it may be more useful to assign the initial identification of relevant legal documents to a research assistant, who may be able to locate many on the internet and download those that appear most important. Documents of general interest such as the constitution and basic organizational laws can be provided to all team members. The remainder can be reserved for review by individuals according to the relevance to the areas in which they will be working.

As noted by one early reader, if team members are illiterate in the local language, translation of legal documents (as well as of other written material) can pose enormous problems.[35] Translations of even common languages such as French, Portuguese, and Spanish often are inaccurate because it is difficult to find translators well versed in the legal terminology of both languages. As the reviewer further noted, "We still often see translation where precise legal terms of art are substituted for better-sounding synonyms that may result in a completely changed meaning." Mistranslations can be comic as well as misleading. One of the handbook authors once had a sentence on "judge-shopping" translated into Spanish as "judge-buying." As the reviewer noted, there is no ideal solution, but at the least, having someone on the team who is fluent in both languages and in legal terminology can serve as a check on the most egregious errors.

C. Informant Interviews

Paper documents can lay still, and the descriptions or evaluations that they offer often are outdated by the time they are released. Although internet publications and websites may be more timely,[36] like their paper cousins, they often lack an empirical basis, represent a biased point of view, or promote an idealized picture of the way things work. Therefore, assessment data collection cannot rely exclusively on the written record. Team members must also engage directly with the people who work in and regularly use the system. Information provided by these people is particularly important in determining actual practice as opposed to what the formal rules dictate. Personal engagement fills also can fill in gaps, especially regarding practices not covered by formal regulations, and help identify the problems encountered by system users (or by those working within the

---

[35] Olga Ruda of ABA/Rule of Law Initiative.

[36] However, websites often are themselves outdated. Organizations may have no systematic program for renewing them, and no one may even notice that information so basic as the names of the members of the Supreme Court is no longer accurate.

system). Observably, there is far greater similarity in the legal frameworks of different countries than in their system outputs. If one wants to know why a country has an unusually high or very low appeals rate, what the law says is only part of the answer. Additional explanations can be sought by asking judges and attorneys why the appeals are made and accepted.

Formal descriptions of duties and functions may not reflect who actually does what or how he or she does it. For example, throughout Latin America, many functions legally reserved for the judge, such as taking depositions (interviewing witnesses), are routinely delegated to courtroom staff. Conversely, prosecutors rely more than the law might suggest on the conclusions reached by the police in their own investigations. Reform team interviews and observation are the only ways to determine the extent to which these departures from law happen and why.

Informant interviews also can be used to collect information on the perspectives, motives, and tactics of individual actors (lawyers, judges, court staff) as they shape system outcomes. The primary purpose of the informant interview is to capture the interviewee's knowledge about system operations. However, the interviewer also should be attuned to what is not said, to what appear to be incorrect statements, or to the occasional offhand remark that can provide a whole new set of insights.

For example, taking two real cases, a court human resources director who does not know how many judges are employed by the court or who guesses wildly on the number, or a head of public defence who does not how many cases his defenders normally carry each indicate critical problems. Interviewers also should be aware of respondents' occasional inclination to provide misinformation to settle a grudge or make a broader point. Finally, as any seasoned interviewer knows, people's memories are not infallible, and they may volunteer inaccurate information only because they do not remember what actually happened. The last is particularly likely when they are asked to provide quantitative data, for example, the "average time" to reach a judgment or how often trial court rulings are overturned. Unless respondents are amateur statisticians, it is highly unlikely that they will manage these statistics well.

Several important considerations arise in the context of interview methods.

1. Selection of interviewees. As their name suggests, "informant interviews" are intended to provide information. Therefore interviewees will be selected on the basis of their ability to do so. This criterion contrasts with the requirements for surveys, which via random sampling and large numbers of respondents, aim at tapping the average or representative response. Which informants are chosen and how many interviews are done depend on what the interviewer wants to know. As a general rule, in informant interviewing, *quality trumps quantity*. Nevertheless, to control for individual biases and gaps in knowledge, even on fairly similar points, it is best to include several rather than a single informant. Again, how many depends on the resources that are available.

As a rule of thumb, the process might start with three interviewees in each major category, for example, trial judges, their support staff, and attorneys handling commercial cases. The interviewer could expand the number if there appears to be considerable inconsistency in the answers. Of course, if the focus is on one agency, many more respondents should be chosen to represent it. Similarly, if broad geographic coverage is desired, the numbers also increase. Inevitably, resource restrictions mean that most interviews will take place in one location, most probably the capital city. Nevertheless, experience suggests that the views in the capital may be very different from those "in the interior." ABA/Rule of Law Initiative relies largely on informant interviews for the numerous assessments it has done for a variety of donor agencies. It reports that it aims for 35–40 key informants, half of whom are from the principal agency (judiciary, prosecution, bar) under study.[37] It also tries to schedule a few interviews in locations other than the capital.[38]

2. Diversification. Interviewees can be expected to project a point of view bound by their roles in the system. A lawyer may give a very different view of delay or corruption than would a judge or registrar. A human rights advocate may emphasize certain symptoms or problems that a business person would ignore. High-level officials may speak from a macro point of view without much knowledge of how the work on the ground is actually done in reality. . In contrast, lower-level actors, system users, and would-be users may have a rich understanding of realities but no overview of the system. Accordingly, *diversification of sources* is an essential criterion in the selection of interviewees. Although system users may be a less valuable source of technical details, including them among the interviewees is important to ensure that their experiences are covered.

Another caveat provided by the ABA Rule of law Initiative involves the insistence on the part of upper-level officials that they provide or at least review the list of informants. Olga Ruda, a staff member, reports that this is a problem in both Eastern Europe and the Middle East. She adds that, in a few cases, it has led to decisions not to publish the final reports because of doubts about the quality of the information derived from the respondents who were willing to talk. While such top-down supervision is not common to all countries, it obviously dampens the interviews. This effect may occur even in countries in which the top leadership appears oblivious. For example, Brazilian researchers, who have relied extensively on surveys of judges and other sector employees for their work, now report that potential respondents are reluctant to participate, fearing possible repercussions for their careers. These examples are not an argument to suppress interviews. They simply highlight the importance of using several methodologies and not relying only on what people are willing to say.

3. Structured vs. unstructured format. Interviews can be organized like conversations, taking the discussion where the interactive process leads; or they can have a structured

---

[37]The ABA Rule of Law Initiative is a public service division of the American Bar Association. It is dedicated to promoting the rule of law around the world by supporting the legal reform process in over 40 countries in Africa, Asia, Central and Eastern Europe, Eurasia, Latin America and the Caribbean, and Middle East and North Africa. www.abarol.org.

[38] Olga Ruda, private communication.

format akin to that of surveys. It may be useful to incorporate both methods, using an unstructured format for much of the discussion but adding a series of questions to be asked of every interviewee. As mentioned, those interviewed will not be randomly selected and they inevitably will be few in number. Consequently, the answers to the standard questions cannot be considered to have the statistical significance of survey data. Nonetheless, it still will be important, for example, that all or 9 of 10 judges interviewed held a similar view about lawyers' tendencies to indulge in frivolous or abusive pleadings to postpone a final judgment or that the judges' tended to admit these motions out of a fear that doing otherwise would only buy them trouble with those overseeing their careers.

4. Corroboration of other sources (triangulation). Beyond diversification, interviews should seek to test one input against another to corroborate observations given from different (or relatively similar) points of view. This practice is known as triangulation. Corroboration involves a check on past interviews by recycling observations into the basic questions posed to newer interviewees. Interviewers should note discrepancies or contradictions, and if necessary, do follow-up or additional interviews as a further check. It is important to recognize that such discrepancies, especially when they refer to issues like usual practices or interpretations of legal requirements, may never be resolved. In this case, their existence itself is an important finding. Here, triangulation, rather than correcting erroneous statements, validates the contradictory views and, possibly, practices.

5. Establishing trust. The quality of information will depend heavily on the interviewees' trust in the interviewer. Justifiably, local actors may worry that candid assessments with attribution will subject them to punitive action. Thus, interviewers should ask whether the interviewee will allow attribution, and if not, ensure confidentiality. Depending on the agreement reached with each interviewee, confidentiality might mean:

- Not including the interviewee's name in the list of sources
- Listing the name but not attributing any remarks in the text to him or her (instead, citing a confidential interview with a court employee, private attorney, while ensuring that specific comments cannot be linked to the anonymous source)
- Listing the name, but attributing only remarks for which the informant is willing to take credit.

To guard against reporting a series of unsubstantiated, hence controversial, claims, triangulation can be used to ensure that confidential statements made by one informant are supported by other sources.

A related theme involves the number of team participants in each interview. Practice varies, from sending the entire team to each interview to sending only one member. This theme is included under trust because of the authors' impression that sending a squad to conduct an interview with one person is not a confidence-builder. Because interviewing is exhausting work and because of the potential for misinterpreting responses, two

members might be ideal, possibly three if one of them has to translate. However, more than three interviewers may make the interviewee feel as though s/he is facing the inquisition.

6. Taking notes. Note-taking is very important. However, because developing trust is so essential to the candor of the interviewee, interviewers may have to choose to refrain from taking careful notes on a notepad or computer or directly recording interviews, even though such techniques are more likely to ensure accuracy.[39] Reporters must exercise their judgment in making these choices. However, it is advisable not to rely on the reporters' memories. Therefore, when notes are not taken during the interview, it is important to write them up immediately afterward.

D. Direct observation

As an important check on the limits of desk and computer research, legal analysis, interviews, surveys, and statistical studies, *direct* observation is a critical diagnostic tool. As one primary example, the observation of judicial processes helps diagnosticians understand and challenge alternative sources of information. In an illustration from an Eastern European country, the interviewer learned that a 3-judge panel hears 70 cases each day. However, direct observation showed that most of these hearings were announced, then quickly deferred due to the failure of critical actors to appear or submit what was required. In another illustration, by observing courtroom practices in action, team members often can identify informal routines that are never mentioned in a law or operations manual but that cause additional delays or provide opportunities for corruption.

The frequency and length of direct observation will enhance the quality of the findings. One visit to a court session is better than none but hardly sufficient as an empirical basis for evaluating court proceedings. Beyond court proceedings, it is also important for team members to observe other critical processes. These include how a case proceeds through the system from filing to enforcement; how notification is done; how the police and prosecution coordinate during an investigation; and how the body that selects judges or prosecutors operates.

Observation is also part of the interview process, especially if interviews are held in the informant's place of work.[40] For example, while waiting to interview a member of a judicial selection body in a Latin American country, the interviewer noted the number of young lawyers waiting, files in hand, to speak with the council member. The obvious suspicion—that they were doing their own lobbying—was largely confirmed by subsequent interviews. However, the issue might never have emerged had the first observation not been made. A casual conversation with an attorney sitting outside the

---

[39] According to Ruda, ABA/CEELI does not tape-record interviews as a matter of principle. However, a team member always can ask the interviewee whether the interview may be recorded.

[40] Ruda notes that interviewing at the work place is an ABA/CEELI preference. However, outside Eastern Europe, many interviewers have found that they get franker responses if the interview is conducted away from the place of business.

office of an appellate judge in the same country revealed that she was waiting to "remind the judge that she liked to win her cases." This chance encounter encouraged further work on the rules (or lack thereof) on lawyers' contacts with judges. In Cambodia, interviews conducted with a public defender outside a prison produced an opportunity to ask about defense strategies. The interview revealed that, in that country, the best that a defender could hope for would be to get his client sentenced for time already served. At that time, the late 1990s, pretrial release or acquittal was out of the question. In this example, the question might have been asked elsewhere, but the location and the fact that the defender was visiting a pretrial detainee provided the inspiration that otherwise might not have occurred.

Less Conventional Quantitative Tools

E. Surveys

Surveys provide a powerful source of information. They are useful for gathering confidential inputs from personnel and users about their experiences and perceptions. Because surveys reach large numbers of informants, they have a scientific weight that interviews rarely can claim. Surveys' structured format does not provide the wealth of detail available from informant interviews. However, unlike the latter, survey results can be more easily aggregated and then quantitatively evaluated, enabling comparisons across time (longitudinal studies), and within and among national systems (latitudinal studies).

In survey design and interpretation, several caveats should be kept in mind. First, because a survey's power hinges on its representative nature, *team members must take care in defining the relevant population and selecting the respondents within it.* If the survey is to be valid, selection must be randomized. Otherwise, the survey has asked the same question of 30 or 3,000 people who may or may not represent the defined universe. Identifying the universe from which the sample will be drawn also can pose problems. Sampling "court users" by taking every tenth person who exits a courthouse is a common technique and looks straightforward. However, it overlooks that (1) these people probably will be largely lawyers, not their clients; and (2) in a country with internet filing, a sample selected in this fashion may exclude the lawyers who use that method.[41]

The second caveat is that *the questionnaire should be constructed carefully*, as what is asked can shape the responses in undesirable ways. Those constructing the questionnaire should check to ensure the respondents understand each question to mean the same as those asking it. If they do not, problems of validity and reliability of findings may emerge. A World Bank survey asking lawyers to estimate the number of steps in a simple debt collection process seems to generate a range not consistent with the relatively

---

[41] This reverses the now-famous error made by public opinion surveys in the U.S that equated their universe of voters as those having telephones (and so drew their random sample from that group.) Because only the wealthier population had phones, the pollsters determined that Alf Landon would win the presidential election. A third possibility, likely in countries that have provided kiosks in courthouses for checking on case status, is that the sample would tap into the hoard of legal assistants sent by their firms to collect information for the lawyers.

uniform legal framework, possibly because the concept of "steps" is itself subject to varying interpretations. Many user claims about court corruption tapped by other surveys appear to originate in a misunderstanding of legitimate filing fees, or in their attorneys' misrepresentation of where their own fees go. Even a question so seemingly simple as whether the courts "obstruct" business operations requires a contextual interpretation. Where businesses do not use the courts because they are very unreliable, entrepreneurs are likely to report a lack of obstruction. Aside from using or seeking advice from experts experienced in questionnaire construction, pretesting with a mixed group of likely respondents can help avoid later problems. Warning signs include little variation in responses (everyone answering the same way could indicate that it is simply not a very good question); other unanticipated patterns in responses (which might reflect reality, but which also may indicate the question is not being interpreted as the designers thought), or respondents' obvious difficulty in producing an answer (meaning they may not understand what is meant).

The third caveat is that *survey data should not be equated with data on real operations*, as surveys usually are based on perception. What actors or users perceive may differ greatly from reality.[42] Delays or corruption may be exaggerated (or underestimated) in survey data. Respondents may have had little or no direct contact with the organizations they are evaluating or may not understand what they are supposed to do. Nonetheless, perception gives important clues about reality, and also has a direct (at times, circular) impact on it. If, for example, a lawyer *believes* corruption to be pervasive (even if it is not), s/he might be more likely to offer a bribe (thus making corruption more pervasive).

---

[42] While some early readers contested this claim, we refer them to Herbert Kritzer (1999, 1983), who has written extensively on the use and misuse of surveys to evaluate judicial performance.

It is unlikely that present-day assessments will have the resources to conduct national-level surveys or even a reasonably representative subnational version. There are exceptions (box 5), but there are also alternatives. One option is the use of surveys that already have been conducted and contain information relevant to the assessment. A second alternative is to purchase questions in a periodically conducted survey. The advantage is the ability to access a rigorously selected sample population rather than having to construct one from scratch. Third, the team member can draw a sample from a much smaller universe (court users in one district) or simply ask a series of questions of a group of nonrandomly selected respondents (possibly, as above, as part of less structured interviews). The results will not have the significance of a well-designed survey, but the fact that 60 percent of the small sample, or 28 of 30 lawyers interviewed, mentioned corruption as a major problem can still be important.

Box 5. United Nations assessment of Nigerian justice: Survey-based approach

There are inevitably exceptions to any rule, and this study offers an exception to the claim that assessments will not have sufficient funding to do their own surveys. This UN assessment was in effect almost entirely survey based. It drew on interviews with 5,766 judges, courtroom staff, court users, and lawyers in 3 Nigerian states. The assessment provides a brief legal and historical overview of the country's justice system and references a case audit, without providing any figures on the results. However, the bulk of its content revolves around the responses to questions asked about access to justice, timeliness of justice delivery, quality of justice delivery, judicial independence, impartiality and fairness, trust in the courts, and corruption. The assessment did not provide information on costs and time, so the reader has no idea how these compare to a more conventional approach.

The product demonstrates both the strengths and weaknesses of the methodology. The problem definition is quite clear, *grosso modo*, and as differentiated by types of respondents. Judges, for example, were less critical of corruption and delays than were court users. Differences among the three states also were noteworthy. However, the analysis of causes and the identification of remedies also were drawn from survey responses, leading to suggestions that, technically speaking, appear inappropriate. Furthermore, there is no accompanying analysis of the system itself. Thus, the local respondents might not have noticed the contributing factors, because they were assumed to be "normal." Examples of such factors are the legal framework, system for selecting judges and staff, and their terms of service, contributions of the private bar, social practices, and conventions. Because the issue of access was posed only to those who already used the courts, the study did not cover obstacles that might prevent getting there, conceivably a far larger problem. Nor did it look at alternative mechanisms, such as customary dispute resolution mechanisms.

Some of these shortcomings might have been addressed by a larger and differently organized survey. However, a better alternative might have been to combine a less ambitious survey with other methodologies, such as interviews, a more extensive case audit, legal analysis, and an institutional review.

In summary, the UN-Nigeria approach was excellent for defining the complaints of those who worked in or used the courts. However, it did not tap the views of those who could not get there, and its reliance on public opinion to derive solutions is highly questionable.

Source: UNODC 2006

It also bears noting that the more traditional attitudinal or perception survey has been joined by survey instruments attempting to tap real behaviors. Called "experiential surveys" by some of their authors, these instruments ask the interviewees such questions as how long their cases took to process, whether they had to pay bribes, and how much they invested in legal fees. While generally regarded as a closer approximation to reality than the attitudinal variation, they are subject to their own caveats. These arise in the following considerations: people's memories often are inaccurate; their understandings of events may be in error (was it a bribe or an extra fee the lawyer pocketed?); and even when guaranteed anonymity, they may be unwilling to report certain events. Nonetheless, even with these shortcomings, the innovations can be useful to understanding the system-user interface.

Another mechanism, reportedly adopted by DFID in its access to justice programs, is the "naming, blaming, and claiming" framework developed in the 1970s by American

political scientists (Grossman and others 1982). This framework attempts to determine when and why people take complaints to the formal authorities. Because it requires lengthy interviews and thus is a costly approach, it is probably not appropriate for an initial diagnostic. Applied later, it can generate useful background for any access program to determine the extent to which poor people in particular have problems that are not well attended by available mechanisms. The framework also has been used (Hendley and others 1999) to understand court use or non-use by firms. As it relies on interviews, even when applied as a survey, the framework could also be incorporated in interviews conducted more subjectively. The results will not be "scientifically sound," but they could still provide useful insights to the team in understanding patterns of court use.

F. Aggregate statistics

1. **Sources.** An important complement to information derived from informant interviews and surveys is the collection of hard data (quantifications of phenomena that can be counted).[43] Examples are:

- Numbers of filings and dispositions, at the trial and appellate stage
- Average times to disposition for different types of cases and the relative frequency of different types of dispositions (for example, by judgment, dismissal, withdrawal of the compliant)
- Ratios of judges, police, or prosecutors to population
- Numbers of prosecutions for corruption
- Salary levels for state personnel.

These data usually are presented in the form of "aggregate" statistics. The term aggregate is used because they provide information on entire collectivities (of cases, employees). As discussed below, this format somewhat limits the types of analyses they allow, especially regarding crossing variables, to determine, for example, which types of cases have the highest resolution rates. However, even the grossest aggregations usually permit some further analysis, and more sophisticated categorization can allow much more.

The situation is gradually improving, but in many countries aggregate statistics are still difficult to gather. Often, there is no one centralized place in which they reside, for the sector as a whole or even within each organization. Databases may be scattered among different offices. Moreover, there may be visible inconsistencies in their counting of similar events. For instance, statistics on homicides or criminal cases kept by the courts, prosecutors, or the national statistics bureau often vary tremendously because of different collection methodologies, definitions, and end uses. In extreme examples, countries have

---

[43] See Legal and Judicial Capacity Building Project (Case Management and Court Administration): Diagnostic Report (2002) (observing that "[m]any of the earlier consultant's reports dealing with the District Courts appear to have a propensity to describe tragedy, possibly due to their extensive reliance on anecdotal evidence rather than hard numbers of actual court performance" and transforming the description of case management and court administration problems "from a collection of anecdotes into hard data collected from case files.")

hardly any quantification of institutional data available in any location. Here the team may have to construct its own statistics, using whatever written records are available (lists of employees or courts; courtroom records of individual case filings) as the basis for rudimentary quantification. Because this task is extremely time consuming, the team must be very selective as to what it decides to count on its own.

Even where statistics are centrally collected, their reliability (and validity) is often problematic. People will manipulate statistics when they realize that the latter can be used against them. Once court personnel realize that they are being evaluated for their productivity, they may begin to fudge the numbers. This manipulation may have occurred earlier or been inspired by the impending assessment. Those using the statistics (including the assessment team) need to be aware of the possibility that the numbers were manipulated, because if this has happened, it would have serious implications for the validity of their analysis.

There is now sufficient evidence of a certain amount of tampering with court statistics to drive the point home. For example, interviews with one Latin American judicial council revealed that judges submitted differing sets of statistics on caseloads to different offices. One set, including backlog, went to the office that determined whether additional judges were needed. Another set, without backlog, went to the office that evaluated the judges' own productivity. Another example found in two other Latin American systems was judges' tendency to refuse to admit cases for flaws of presentation, thereby reducing their workload and upping their disposition rate. Sometimes entire judiciaries indulge in such practices, as when they are asked to provide productivity statistics to an international database.

Figures on budgets and salaries also are problematic. Countries commonly include varying mixes of agencies in the "judicial budget." Furthermore, reported salaries often do not include bonuses or other payments on top of the base amount. In some African and Asian countries, housing, some security, and sometimes transport are provided to the judges but may not be reported as part of "salaries." In other countries, such as Brazil, the judiciary handles funds set aside by the government to make payments to parties that win cases against it. The team must take care that these funds are not included in the overall budget figures.

World Bank and other researchers working with sector statistics frequently have found indications of inconsistent classifications within a single country, abrupt changes from one year to another in how things are counted, and inaccurate or incomplete recording.[44] Because recording statistics often appears to be a pointless, extra burden to those tasked with doing it, the job often is assigned to the most junior members of a judge's or prosecutor's staff. The errors cannot be corrected after the fact, but a good statistician may be able to work around some of them, or the numbers can be cited with all of the appropriate cautions as to their ultimate reliability.

---

[44] See World Bank 2005 for a discussion.

2. Analysis. Once the team has collected aggregate statistics, the question becomes how to use them. There are two basic answers. Statistics can be used to check and expand on descriptive statements garnered through more qualitative approaches, and they can be subjected to further analysis to identify patterns and tendencies in national operations.

Twenty years ago, such data, when they were even available, stood on their own and were hard to interpret. In the absence of a comparative yardstick, knowing that a judiciary receives 6 percent of the national budget or that there are 100 police officers per 100,000 inhabitants is not terribly helpful.[45] Today, a growing number of international, regional, and subregional data sets provide the bases for comparative analysis. They enable a team to determine, for example, whether complaints about delays or excessive workload are borne out by the facts.[46]

Of course, data also must be interpreted contextually. A "low salary" for judges or prosecutors must be assessed not only against international levels, but also against salaries for comparable positions within the country in question and against the local cost of living. An average caseload of 400 annual filings per judge might be high or low depending on the nature of the cases, procedural rules, and prevailing expectations as to individualized treatment.[47] It should be recognized that there is no absolute criterion regarding target salaries, caseloads, or ratios of employees to population. However, there is now an idea of the reasonable range of variation. This range provides a basis on which to determine when figures are sufficiently unusual to suggest performance problems.

Aggregate statistics also lend themselves to other kinds of analysis. In some countries, data are available for several years. If so, assuming that the subcategories have not disappeared in the final aggregation, changes in overall caseloads and variations among types of cases or courts can be identified. In some countries, several types of aggregate statistics are available, including numbers of employees, number of filings and dispositions; number of cases entering at the trial and appellate level, reported crimes, criminal investigations, and criminal cases processed. In this circumstance, simple calculations can be used to determine average caseloads, clearance rates (dispositions over new entries), rate of backlog accumulation, appeals rates (ratio of appeals to first-instance judgments), and responsiveness to exogenous events such as changes in crime levels. Sector data can be crossed with population and economic statistics, changes to basic laws, or other historical events to track national and subnational trends in service distribution and the impact of exogenous factors on overall demand. If the statistics are

---

[45] For the reader's information, 6% is very high, and 100 per 100,000 is quite low by international standards. However, theses percentages may be appropriate for the countries in question. The percentage of the budget has no meaning if the team does not know what the budgets include. In a social democracy with high levels of social security, public health care and education, the percentage dedicated to the judiciary is probably no more than 0.5 percent. However, in absolute terms it may well be enough to sustain a high-quality organization.

[46] Useful sources include Blank and others 2004; CEJA 2003 and 2005; and CEPEJ 2005 and 2006.

[47] It might be mentioned here that, in Europe, debt collection cases usually are rapidly dispatched because for the most part, they are not regarded as real controversies. In Latin America, they often become mini-trials, thus requiring more judicial input. In the same vein, Brazilian courts batch process complaints about improper adjustments to public pensions, turning out thousands of decisions in a single month.

disaggregated by work unit (court, prosecutorial office, police station) or district, or by type of case, it is possible to track within-country or within-system variations. Possible findings of relevance to diagnose system performance and needs include:

- Radical increases in demand, due to economic or political events or legal change. An example of the latter would be the addition of procedures intended to expand access to certain population groups);
- Unequal distribution of service units (either against population or workload);
- Tendencies for certain kinds of cases to be processed far more slowly than others.

Figures 1 and 2 provide examples of the kinds of simple analysis that can be done using aggregate statistics.

Figure 1. Growth in caseload in Brazilian courts. First-instance filings.
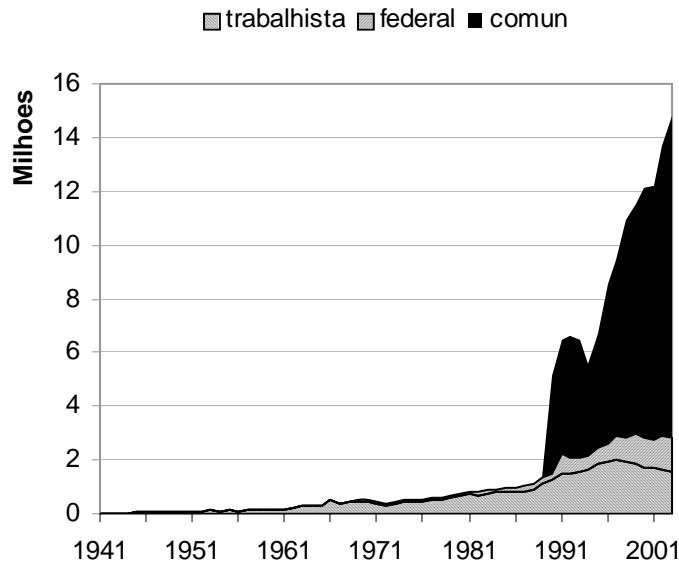
Source: World Bank 2005.
Note: Labor Courts 1941–- 2001 - 2003; Federal Courts 1967–2003; State Courts 1990–2003.

Chapter 5 provides more information on caseloads. It also explains some basic operations that can be carried out with simple statistical figures in order to learn about possible problems in a court system or in a single court.

Figure 2. Average caseload per judge across Brazilian court systems1999–2003

Legend: trabalhista ◆ — federal □ — comun △

Source: World Bank 2005.
Note: Brazil annual first instance dispositions per judge (1999–2003) in two national court systems and the state courts: Labor (trabalhista), ordinary (federal), and state (comun).
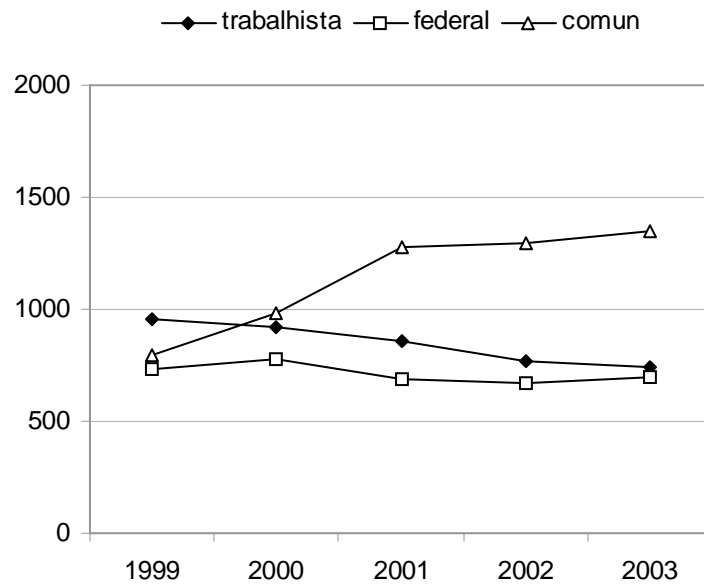
G. Case File Analysis

As suggested above, statistical databases need not be treated statically. They can be used to generate additional information on system characteristics, track chronological trends, and identify performance variations within a national system. Nevertheless, because of their aggregate nature, they are most useful for showing the global picture and less useful as a means of tracking variations in how cases are handled or at what points they encounter bottlenecks. Thus, these databases can be usefully complemented with other methodologies, for example, analysis of random samples of case files, which may be followed from filing through all instances to final decision and enforcement.[48] Although most commonly applied to courts, case file analysis can also be used to track the output of other institutions that keep written records of their cases, that is, police, prosecution, and defense.

As courts begin to improve their internal organization, the case file analysis method is increasingly feasible. However, there may be problems in getting access (illegal in some countries for anyone but the parties and the judge), physically locating the files (where archiving systems are chaotic), and designing a randomized sample (especially if courtroom records are not well kept). Moreover, the exercise is costly and time consuming, and for most assessments could at best be applied to a small group of cases in

---

[48] For examples, see World Bank 2002, and 2003 a and b.

a court or two.[49] As with observation of courtroom practices, one example may be better than none. Nonetheless, unless a random sample is used, with additional qualifications concerning from where or whom it is drawn (that is, nationally or in a single district or courtroom), more limited applications should be understood to be heuristic at best. Additionally, if the analysis will be limited to only a few cases, the team should take care to select "more representative" cases. It is always tempting to select the worst case, such as the one that went on for 17 years without resolution, or in which the various decisions clearly contradicted usual interpretations of the law. However, to understand how the system *usually* works, these are the least useful examples.

Whether working with a randomized sample or a few cases, findings can often be converted to flow charts, showing case trajectories and the most common outcomes. Figure 3, taken from a World Bank study of debt collection cases, is one example of this methodology. Using a sample of 400 cases from 1 judicial district, it demonstrates what happens to those that did not reach sentencing (via what are called here formal or informal exits). These distinctions would not have been visible using aggregate statistics. Furthermore, the latter consider cases exiting "informally" to still be part of the court backlog.

---

[49] Countries such as Australia, Canada, or the United States, which have done this on a national basis, have spent millions of dollars on the process, as well as invested several years in its completion. See Australian Law Reform Commission (2000).

Figure 3. Mexico debt collection cases: Formal and informal exits in three moments of the proceedings

Out of 100 Admitted	**Embargo & Notification 38%**

**FORMAL EXIT
19%**

Dismissal	12%
Formal Withdrawal	66%
Payment	4%
Settlement	18%

**INFORMAL EXIT
43%**

De fact withdrawal	56%
Exhortos	25%
Debtor refuses notice	3%
Debtor cannot be found	6%

Out of 100 Notified	**Judgments 54%**

**FORMAL EXIT
15%**

Formal Withdrawal	36%
Payment	43%
Settlement	21%

**INFORMAL EXIT
31%**

Informal withdrawal or lack of further action

Of 100 judgments that favor the plaintiff	**Auction or Appropriation 3.6%**

**FORMAL EXIT
23.3%**

Payment	35%
Settlement	15%
Plaintiff desists	50%

**INFORMAL EXIT
76%**

No information

Source: World Bank 2002.
Note: "Embargo" refers to the attachment or seizure of assets. "Exhorto" is the request that a judge in a different jurisdiction take an action needed to proceed with the trial.

Case file analysis (and the next technique, evaluation of decisions) naturally requires facility in the local language and knowledge of the legal rules and processes. Thus, this method is one in which the participation of local actors is absolutely essential. WB researchers using the technique have commonly contracted local teams to do the data collection and initial analysis. In a participatory assessment, local counterparts may simply volunteer staff to help out. Aside from the knowledge generated, the advantage for them is the ability to learn a new methodology that they can apply later on their own. Although cost and time constraints may preclude using this methodology in the initial assessment, it is highly recommended that it be conducted in the later project preparation or early stages of implementation. There are some questions, such as those about barriers to access, for which case file analysis is less appropriate. However, where the questions revolve around delay, predictability of judgments, or even biases and corruption, it is an essential tool.

Less Conventional Subjective Methods for Collecting and Analyzing Data

While quantitative methodologies represent one recent addition to assessment techniques, they are not the only ones. Those doing assessments also have begun to expand their soft or subjective repertoire, adding new types of data and different kinds of analysis. Examples follow.

H. Evaluating the Quality of Judicial Decisions

Beyond the examination of inputs and case trajectories, judicial decisions provide an essential source of information to evaluate. The results reached, accuracy and consistency of legal analysis, thoroughness of factual summaries, and quality of the rationales underlying the decisions each provide important elements to examine. As one of the softest (most subjective) methodologies, this approach has many detractors and little guidance as to how it is best done. It also is a technique that external donors may prefer not to use because of the risks involved. However, local counterparts often are quite insistent on its inclusion although their recommended criteria may not strike external participants as terribly appropriate. An example would be the number of constitutional citations in first instance judgments on debt collection cases. Perhaps the solomonic solution is to leave the exercise to the counterparts so that they can dispute the methods and the results.

As with case file analysis, with which it can be combined, there is always the danger of selecting a few outrageous, and therefore nonrepresentative, cases and overgeneralizing from the findings. Where statistics permit, some quantitative mechanisms can be substituted, for example, comparing pro-defendant or pro-plaintiff decisions in a number of similar cases to evaluate if not the quality, than at least the uniformity of decisions. If information on the parties is available, such techniques also can be used to investigate

alleged biases.[50] Similar techniques, with all the same caveats, also can be applied to nonjudicial decision-making (for example, ADR or traditional dispute resolution services) or to police, defense, and prosecutorial operations. The additional problems here are that records may not be kept or that access to those that are may be highly restricted.

I. Incentive Analysis

Similar to analysis of judgments, incentive analysis is new but very subjective.[51] It also is less a source of data than a means of processing those collected through other methods. However, incentive analysis increasingly is recognized as an important contribution to any assessment because of the realization that many problems (and impediments to reform) are the function of an adverse incentive structure. By incentives, this handbook means the factors in addition to the official rules that determine how actors perform their functions. Examples might be why summons servers might not make an effort to find the affected party (insufficient resources, bribes, or just a lack of supervision?), or why judges might allow frivolous complaints (fear of repercussions from those placing them, no effect on their own evaluations, bribes?). As with the evaluation of judgments, understanding incentives is an art rather than a science. Nevertheless, it still merits adoption because of its criticality to understanding why a system operates as it does. Why, for instance, do similar rules produce a 3 percent appeals rate in one country and 30 percent in another? Why are judgments never enforced in one country but have a relatively decent enforcement rate in a second? Why is reducing delay not a self-evident goal for court personnel?

To ferret out the hidden incentives, team members can apply certain rules of thumb that have proven useful in a variety of settings. For example:

- Knowing how much different categories of actors are paid and how pay is linked to performance can provide insights into how they operate. If lawyers are paid by the procedural step, they will operate differently than if they receive a portion of any final award.
- Systems for monitoring and evaluating performance can change performance incentives, but their efficacy depends on whether those doing the monitoring have any incentive of their own for doing it. Judges may theoretically have responsibility for overseeing courtroom staff, but if they themselves do not suffer any consequences from low output, it is unlikely that they will spontaneously exercise any real control.

---

[50] For example, in ESW done in Mexico in 2000 (World Bank 2002), the researchers investigated the alleged pro-debtor bias in debt collection cases. Their findings, based on the direction of the judgments, suggested that the bias did not exist.

[51] Again this statement met opposition from some early readers, but given the need to "get into the head" of the organizational actor, it is hard to see this as anything but subjective, unless one takes the rational-choice position that it all can be reduced to monetary concerns. The authors believe that other incentives (reputation, family responsibilities) matter, recognize that risk tolerance varies among individuals, and accept the other findings of behavioral economists (Sunstein 2000). Consequently, the authors think that reducing it all to money is a nonproductive oversimplification.

- Knowing who makes decisions on appointments, salaries, promotions, and credentialing (as for private attorneys) and their own incentive structure is also critical. The latter depends on who appoints them and why. If they are political appointees, they may be expected to represent their parties. If they are appointed from within the institution, different kinds of incentives would apply.
- Knowing about budgets and their allocation helps to understand why there is often not much enthusiasm for improving efficiency because the results will not benefit those who have to implement the changes.
- Transparency (information supplied to outsiders on institutional workings) also can be a factor. Generally, greater transparency will encourage more conformity with official rules. However, it also can lead to endless bureaucratic paralysis.

Although incentive analysis logically follows the collection of other data, it will be important to plan for its conduct throughout the rest of the assessment. Otherwise, one may not have the necessary information. As they proceed with the rest of the assessment, team members may want to collectively discuss the key actors in the processes studied, the emerging hypotheses as to how they affect outcomes, and the factors that may shape their actions. Even if one team member eventually does the analysis, ideally the inputs come from the entire team.

J. Focus and Study Groups

Focus groups have their own detailed methodology, but their proponents believe that they often are a better means of eliciting information than surveys or informant interviews. The difference lies in the group dynamic and its ability to lead discussion into areas the outside survey designer or interviewer might never discover. Whether focus group results constitute hard or soft data is still a matter for debate, but it is evident that the groups can be a productive source of insights not attainable through other means. Assessment also can be aided through the establishment of study groups of diverse expertise and representing different constituencies in the system, regional counterparts, and expert and donor communities. The institutionalization of local teams maximizes expertise, political support, the effectiveness of dissemination strategies, and the cultivation of actors to carry out future recommendations developed in the diagnosis. The integration of national and comparative law experts and donor representatives further strengthens teamwork and collaboration toward shared ends.

K. Preliminary Reporting Feedback and Collective Interactions (Workshops, Conferences, and Seminars)

The report serves as both product and process. It captures what has been understood and presents itself for further review, criticism, feedback, and input. The development of a diagnosis is dependent on the investment in this interactive process. Accordingly, sharing drafts with individuals, small groups, and larger audiences is critical to identify inaccuracies and misunderstandings in advance of the release of a final report. Especially when directed at discussing the findings from the methodological tools discussed above, workshops, conferences, and seminars provide additional means to collectivize important

input into the diagnosis. Local study groups provide a foundation for planning and implementing these events. Reporting the inputs is critical to capture the value of these interactions.

5. COMMON COMPLAINTS

This chapter highlights the three most common complaints about justice systems that come up in assessments. These are excessive delay, insufficient access, and corruption. The most relevant aspects of each complaint for the assessment will be discussed:

- Why they are important.
- How they are defined and any related problems of conceptualization.
- How their presence and dimensions can be determined.
- Where proximate and underlying causes are most often sought.
- What types of remedies have been found useful.

1 Common Complaint: Delay

The most common complaint about justice system performance is excessive delay in processing cases. There is an almost universal opinion that courts and other justice sector institutions take too long to provide responses to users of their services. "Delay" refers to a task being late or deferred, thus exceeding a standard. Another term frequently used in this context is "backlog." "Backlog" means an accumulation of tasks to be handled. If the number of tasks is so great that it interferes with its own completion, the term "congestion" applies.

The International Covenant on Civil and Political Rights (Article 14), and regional human rights treaties specify that cases must be disposed by courts "without undue delay" or "within a reasonable time."[52] Timely justice is not merely an abstract right; delay has multiple broader impacts. For example, when a case goes on for years, suspects, witnesses, or evidence may disappear. In civil cases, delay can discourage legitimate plaintiffs as well as reduce the value of any awards that they eventually receive. In criminal cases, on the one hand, delays can work enormous hardships on those under suspicion or accused of crimes, including, but not limited to, their lengthy pretrial detention. On the other hand, delay can be used to avoid justice by extending investigations and trials beyond the statute of limitations for crimes, or otherwise making their successful pursuit less likely. Delay also obstructs access to justice because it works in the favor of those, usually the better off, who can best tolerate the delay. It also may encourage corruption by offering opportunities to request bribes to speed up case processing (or to hold off attention to a case).

As often is the case with intuitively simple concepts, the challenges posed for those attempting to analyze delay are far more complicated. Relying on popular perceptions can be extremely misleading. Extended experience working in countries in which delay was first identified as a major problem often reveals that its incidence was exaggerated; its causes were not as claimed; and its impacts were different than imagined. Thus, assessments will have to treat the topic with care, starting with an effort to introduce

---

[52] The European Convention for the Protection of Human Rights and Fundamental Freedoms (Article 6), the American Convention on Human Rights (Article 8), and the African Charter on Human and Peoples' Rights (Article 7). See appendix 2 on international instruments that lay down standards.

quantitative parameters. To determine whether delay in fact exists, two steps are needed: (1) measure processing times and (2) evaluate them against standards to decide whether the times indeed are "unreasonable" or excessive. Neither step is easy.

The first step is to measure processing times. Every system has a few notorious cases that have taken extraordinarily long times to be disposed. However, for assessment purposes, these cases are not representative, and they do not reveal the necessary knowledge about practice in the justice sector. Instead, what is needed is a representative figure—the average (mean or median) times required—as well as a measure of dispersion around the mean or median to show the extent of variation.[53] Few justice systems anywhere have maintained statistical records that enable the easy calculation of these figures for specific stages or for the entire trajectory of a case. Because assistance projects are unlikely to operate in such countries, other mechanisms likely will have to be used in developing nations.

The *Doing Business* methodology (World Bank 2006a) asks lawyers to estimate the processing times of commercial contract enforcement cases that include hearing witnesses. The time usually begins at filing and ends at completion of enforcement. In practice, in some countries in Western Europe and in the United States, these cases are estimated to take approximately 1 year (300–400 calendar days). In other countries, similar cases are estimated to take up to 1200–1400 calendar days. In most countries, real statistics for these cases are not readily available. The *Doing Business* approach of using estimates by local lawyers does not produce very accurate measurements. However, it may be the best that the team can manage during the assessment. Random samples of case files also can be drawn, subject to all of the caveats listed concerning case file analysis in chapter 4. There are also proxy indicators, such as disposal, clearance, and congestion rates, which are discussed below. These do not directly measure times to disposition; instead, they focus on agencies' abilities to keep up with their caseloads.

The second step is to determine whether the acquired processing times are excessive or not. There have been recent attempts to set international or national standards. The European Court of Human Rights has ruled that, for criminal cases, two years is a reasonable processing time, counting from the moment the reasonable suspicion arose to the decision in the first instance. The American Bar Association (ABA) has set time disposition standards for general civil cases in US courts: 90 percent of all filed general civil cases should be disposed within 12 months after filing, 98 percent in 18 months, and 100 percent in 24 months (NCSC 2005). Some countries incorporate time limits in their procedural codes. These standards have legal relevance. Nonetheless, for

---

[53] As this is not a handbook on statistics, little more will be said about these concepts. The mean is the arithmetic average, and the median is the figure at the fiftieth percentile, that is, half of the universe is below it and half above. The median generally is preferred to the mean because the former is less distorted by asymmetric extremes (a few very high or very low scores). However, the measure of dispersion also is informative. It makes an enormous difference if most scores are clustered around the mean or median, or if the "average" hides a wide range of scores. For example, there is a telling difference between 90 percent of cases being processed in 6 months versus a 6-month average composed of times ranging from a few days to several years.

assessment purposes, they have limited value because they are fairly general and abstract.

The assessment team thus may find it more useful to compare case processing times to those available in international databases *for comparable types of cases.* If so, in estimating average times within the country in question, it will be best also to distinguish types of cases. The average time for processing any civil case is far less telling than the average time for a bankruptcy case, a simple debt collection proceeding, or a dispute over unfair dismissal. Similarly, on the criminal side, simple misdemeanors should be separated from major felonies, and those in turn from complex investigations of suspected corruption or other white collar crimes. An assessment may not be able to capture this information, but if an eventual reform is to focus on reducing delay, it clearly should incorporate more of this kind of analysis early in the process.

What of situations in which existing statistics are so poorly kept and so inaccurate that neither assessment nor analysis is possible? Alternative means exist to reach tentative conclusions on the existence and prevalence of delay. These means require the presence of certain statistics, not only the kind that would enable the direct measurement of processing times. Most of these alternatives focus on agencies' ability to keep up with caseloads so look at factors such as backlog accumulations and clearance rates. All that is required for this level of analysis are the figures on cases pending, entering, and resolved within a given year.

Clearance or disposition rates are calculated by dividing the dispositions by the new entries. A figure of less than 1.0 means that an agency is not keeping up with its caseload so is accumulating backlog. A figure of over 1.0 means that the agency is both keeping up with its cases and reducing its backlog. The disposition rate does not consider the existing backlog or its impact on court performance.

The *congestion rate* expresses the relation between the total caseload and the cases disposed. The caseload consists of the backlog plus the cases entering in a given period. The congestion rate is calculated by dividing the caseload by the number of cases disposed. A congestion rate of 1 (or 100 percent) means that the court resolves its entire caseload. A congestion rate of more than 1 signals that a backlog gradually is being reduced, maintained, or, as the number gets higher, is building up (Dakolias 1999). Experience demonstrates that congestion and disposition rates tend to vary not only by courts but also by types of cases. Consequently, a more thorough analysis would require breaking down the figures by categories of cases. Again, this possibility often is limited by the way that statistics are kept.

In criminal cases, the rates of pretrial detention have been used as proxy indicators for delay. As table 3 shows, the latter rates vary considerably around the world:

Table 3. Long pretrial custody by country, 2005

Country	Prison inmates untried and unsentenced (%)
England and Wales[1]	34.9
United States [2]	21.2
South Africa [3]	27.0
Nigeria[4]	64.3
Honduras, Paraguay, Uruguay	90.0
Dominican Republic, Haiti, Peru, Venezuela[5]	65.0–85.0
Argentina[6]	89.0

Sources:
1. Population in Custody, Quarterly Brief: April-June 2005 England and Wales.
   www.homeoffice.gov.uk/rds/pdfs05/prisq205.pdf
2. www.prisonstudies.org/
3. September 2006 statistics published by the South African Department of Correctional Services. www.Dcs.gov.za
4. International Centre for Prison Studies at the University of London: Prison Brief for Nigeria.
   www.kcl.ac.uk/depsta/rel/icps/wrldbrief/
5. www.hrw.org/advocacy/prisons/americas.htm
6. *Human Rights Watch: World Report 2006*. www.hrw.org

These figures reflect the work of all institutions in the criminal justice chain: police, prosecution, courts.

The range of institutions involved raises another point, covered in more detail below: the need *not only to identify delay but also to determine at what point in the process it is most likely to occur*. This need is most evident for criminal cases because each stage tends to be controlled by a different organizational actor, but it is true for civil proceedings as well. For example, in a study of debt collection cases in Mexico (World Bank 2002a), the case file analysis methodology revealed that delays were greater in the enforcement than in the prejudgment stage. It also suggested that many prejudgment delays hinged on the work of the bailiffs.

After determining that delay is a significant problem, *the next challenge is to identify its causes.* Whether a more thorough examination of the processes can be undertaken depends on the availability of data from each stage. However, even when obtaining data is not feasible, much of the following discussion is still relevant. A first consideration is the size of the caseload and any recent changes in it. Caseload size is the most common explanation for delay, and as such, it is vastly overrated. Caseload size alone does not determine delay. Only in relation to the human and other resources available to dispose it does caseload have meaning. In general, case influx in the courts is determined by

- Legislation regulating which disputes or matters need to be brought before a court
- Incidence of these matters in society
- A large number of cultural and other factors influencing whether cases or matters actually come before a court (Genn 1999).

These factors all can change. Congestion, backlog, and delay are inherent risks for organizations that face uncontrolled demand, while their resources are limited and inflexible. This kind of disconnect is very common for public service institutions, and the institutions in the justice sector are no exception. In principle, anyone can take a case to

court, and the factors influencing the decision to do so are largely outside the court's control. For instance, the demand for commercial court cases will fluctuate inversely with the economy. If more people and businesses are unable to pay their bills, more claims will be filed with the first instance courts. The demand for criminal court cases may well vary with the political climate and public sector priorities in a country. Criminal case levels are related to economic cycles as well. The actual caseload also is influenced by the types of cases that have to be dealt with. A rise in more complex cases, as when a country decides to prosecute more fraud or corruption, may not be reflected in the numbers, but it will constitute a greater demand on resources. At such points, more detailed information on the types of cases is very helpful. Assuming that any kinds of records are kept, the assessment team members should be able to determine whether the size of demand, or any sudden changes in its dimensions, in relation to the resources available can account for the delays that they have observed.

A second set of explanatory variables can be sought in the legal framework as it affects how the demand, once generated, is processed. Procedural complexity has been identified by several observers as a major source of delay (World Bank 2006a; Botero and others 2003). Moreover, when the law does not give judges much ability to override parties' delaying tactics, or when other factors (expectations, custom, or corruption) discourage judges from applying what power they have, similar legal provisions may be abused to create delays in one country and not in another. Thus, in reviewing legal explanations, the team will have to review both what the law says and how it is applied in practice.

A third set of explanations can be found in *organizational practices and structures not necessarily mandated by law.* Donors' emphasis on improving courtroom administration works on the hypothesis that inefficient practices at this level create their own delays as well as the opportunities to create more. For example, where internal record-keeping does not allow judges to know how individual cases are progressing, staff can take all the time that they wish, often putting difficult assignments at the bottom of the pile. Case files can go missing, or staff can be paid to misplace them. Analysts often find that over time certain redundant steps have been added for no particular reason except perhaps to provide work for equally redundant staff. If staff is not graded on performance, no one will have any reason to work more rapidly. It bears mentioning here that *rapid processing of cases has no intrinsic pay-offs for anyone in the system.* There will always be more work tomorrow, whether it is held over from today or is a new set of demands. Hence, *lack of oversight at all stages of the proceedings* is another contributor to delays.

Finally, explanations can be sought in a series of *perverse incentives*, starting from the proposition that there is always someone who benefits from delay and that unless their ability to set the rhythm of movement is controlled, they will have their way. Defendants have an interest in trying to hold off or prevent decisions against them. Lawyers may simply want to string out cases as a means of increasing their fees. However, for whoever wishes to succeed in causing delay, the system or someone within it will have to accommodate the delay: courtroom staff, judges, prosecutors, witnesses, the opposite party's lawyer One cannot change the motivations of the principals, but one can reduce their ability to get the necessary collaboration from the within-system actors. Finally,

there may be a perverse incentive for courts to keep a backlog. A backlog can be a powerful argument in negotiations for more budget.

In proposing remedies, the process can start with the *four major sources of delay:*

1. *Size of demand,*
2. *The law,*
3. *Organizational structure and practice,*
4. *Points at which perverse incentives can operate.*

2 Common Complaint: Access to Justice (Analyzing Problems of Access)

Equitable access to justice is a widely recognized aim of sector reform, underscored by its mention in international conventions and national constitutions. Although its precise meaning is still debated, it is generally associated with:

1. Equal protection of the law for all citizens and others in the national territory and
2. Equal opportunity to seek and receive remedies for alleged violation of one's legal rights by public or private actors before courts and other conflict resolution mechanisms.

The two parts work together. While a well functioning system should afford all citizens the ability to protest violations, it also should make the protests less necessary through its deterrent effects. Realizing that their actions will be sanctioned, would-be violators refrain from doing what the law prohibits or automatically do what it mandates. Obviously, no system works this well, but the ideal underlies the frequent observation that access to justice should not be equated only with access to courts. In fact, high levels or sudden increases in court use can be either signs of system malfunction or an initial, short-term consequence of more positive systemic change.

Two decades of growing attention to access issues have produced a wealth of information on the barriers affecting the poor and other marginalized groups in particular, and on a variety of ways to overcome these barriers. However, experience has been less helpful in:

1. Developing means to assess and measure impacts, except on the immediate beneficiaries;
2. Comparing and quantifying levels of access within or across countries; or
3. Developing integrated strategies that incorporate the variety of mechanisms that citizens may use to resolve their problems.

Just as sheer increases in users of courts and alternative mechanisms do not necessarily equate with improved access, objective need (the starting point) is hardly self-evident. Low litigation rates may indicate important barriers, the availability of satisfactory alternatives, or a relative absence of justiciable conflicts.

For decades, the high litigation rates in the United States, as compared to many Western European countries, have been attributed to the former's weaker social safety nets, its

more aggressive legal culture, and the structure of legal and court fees (greater use of contingency fees and the tendency to have each party, rather than the loser, pay the associated costs). More refined measurements of access would require better information on who litigates or uses alternative mechanisms and with what results, as well as an understanding of what kinds of conflicts go unaddressed for lack of means to resolve them. The "naming, blaming, and claiming" framework developed by Grossman and others (1984) has been applied to these issues (Genn 1999; Hendley 1999), but it is costly and time consuming. Simpler surveys asking citizens about their likelihood of going to court and the reasons why might address "access to justice" in a narrow sense (who goes to court). However, they do not get at the questions of need, availability of alternative mechanisms, or the effects of the current situation on their well-being. In addition, these surveys are costly if applied to entire national samples and stratified to tap target groups.

Absent this baseline data, most access programs instead start with an intuitive identification of groups who require more or better access and of the barriers to this end. Most programs focus on the poor, women, and marginalized ethnic groups. However**,** some program target entrepreneurs because of the negative impact of unsatisfactory access on their activities. Here "unsatisfactory" refers not to their ability to get to court, as they have the means to do so, but rather to the quality of the response they are likely to receive—or access *to justice* in a more fundamental sense. For both groups, factors impeding and discouraging access usually are sought in the following:

- Weaknesses in the legal framework (different aspects depending on the group in question), including inadequate substantive laws and overly complex procedures
- Slow responses of judicial actors due to court congestion, inadequate supervision, judges' failure to control dilatory practices, or poor work habits
- Perceived or real corruption and biases on the part of judges, their staff, and other sector actors (for example, police).
- Poorly prepared or motivated judges and lawyers
- Political interference in decisions
- Inability to enforce judgments due to judges' lack of power or the malfunctioning of auxiliary agents and agencies.

For the poor, women and marginalized ethnic groups, the above list is usually extended to include:

- Physical inaccessibility of services and costs of traveling to what is available
- Legal and court fees and the shortage of free legal services
- Linguistic barriers
- Distrust or lack of understanding of court and other sector operations
- Inability to protect themselves from negative consequences: reprisals; impacts on employment, relationships, or community standing (reputation as a trouble-maker).

Although programs intended to enhance access for disadvantaged groups often focus on the second list, it has been asked, with some reason, whether insufficient attention to the

common barriers might undercut their benefits. As one observer (Faundez 2006) notes, giving citizens better access to corrupt and inefficient courts may be doing them no favor. Thus, a second pro-poor access strategy introduces alternative services or increases attention to traditional conflict resolution mechanisms. Critics do caution that:

- New or traditional alternatives may suffer the same vices as the formal system.
- Both may result in a second-class, inferior justice for the poor.
- Improving traditional systems may be no easier than reforming the courts themselves.

A third set of "access" programs supports litigation intended to help the poor *access* additional (nonjudicial) rights and services often denied them or to empower them politically. This third set is less about justice sector reform than about using the sector to promote broader societal change.

The four general conclusions to be drawn from this still very incomplete project are that:

1. The choice of feasible mechanisms to enhance access is highly context specific because both the barriers and the impacts depend on the local situation.
2. Our skill in removing barriers has outrun our ability to measure and compare systemic impacts.
3. Our capacity to understand and work with traditional systems needs strengthening.
4. The ends pursued through access programs vary considerably, although all would benefit from a stronger, more empirically based strategy.

As these conclusions suggest, the challenges facing the assessment team in this area are considerable. As with delay (and corruption), measurement is an issue, but it is compounded by the ambiguity of the goals and the difficulties of defining and determining the relative significance of the initial problem. Remedies exist in abundance, but their worth can be calculated only on the basis of the answer to these more fundamental questions.

3 Common Complaint: Corruption
The very common complaint of corruption in the justice system is doubly relevant. Justice sector institutions have international standards prescribing their integrity (appendix 2). They also are the very institutions that are crucial in combating corruption in society.

Forms of corruption vary, but include bribery, extortion, cronyism, nepotism, patronage, graft, and embezzlement. The most common form is that of payment for a favor: access to a service or speeding up processing. Some more formal concepts used to distinguish among types of corruption are:

- *State capture*: Firms shaping and affecting formulation of the rules of the game through private payments to public officials and politicians.

- *Influence:* Doing the same without recourse to payments.
- *Administrative corruption:* "Petty" forms of bribery in connection with the implementation of laws, rules, and regulations (World Bank 2004).

The judiciary should be independent, but others, including other branches of government, may attempt to influence or capture judicial decision-making and judicial appointments. Evidence suggests that while in most developed countries corruption is the exception or an isolated incident, in developing countries, it can be increasingly systemic as institutions become less and less functional (Johnston 2005). It also is important to keep in mind that what is, for instance, helping one's relatives in one context may be considered undesirable corruption in another (Rose-Ackerman 1999).

Measuring corruption in the statistical sense is not a straightforward matter, since the participants generally are not forthcoming about it. Perception and experience surveys are the tools most frequently used to identify corruption. These surveys can provide important insights into corruption in a country. Both experience and perceptions aid in gaining a deeper insight into the incidence and levels of corruption in justice sector institutions as well as into the linkages between causes and consequences. However, they are hardly foolproof. Users' perceptions of corrupt practices in the justice system in some cases may be due to delays or incompetence, or to **a** general feeling among the population that all public servants are corrupt. Actual experience may not always be truthfully reported because respondents are reluctant or uncomfortable to admit that they have paid bribes. Justice officials may be reluctant to report corruption in their own profession or peer group.

Some of the leading academic approaches to corruption focus on power and checks on that power (Klitgaard 1999) and on incentives and risks (Rose-Ackerman 1999).
In the power and checks approach, judges or other justice sector actors have

- A monopoly over legal dispute resolution, for example, when there are no viable alternatives such as arbitration or mediation
- Broad discretion, for example, when judges are considered to be independent from review or when prosecutors have wide discretion in deciding whom to prosecute
- Limited accountability, for example, in non-oral processes, nonpublic venues, noncontinuous trials, nonparticipatory dispute resolution, or unpublished decisions, which would be far more likely to be corruptible and corrupted.

This focus on power may explain the conditions conducive to judicial corruption. It does not begin to grasp the motivations underlying corruption in the judicial process

In the approach focusing on incentives and risks, the main determinants of corruption are located:

- At the level of available benefits
- In the riskiness of corrupt deals

- In the relative bargaining power of briber and bribee.

This second approach identifies critical motivations that are applicable to the judicial process. In light of poor salaries, difficult working conditions, and scant resources, judges in many countries have strong incentives to take bribes or gifts. The chances of being caught, disciplined, or prosecuted, as well as the nature and size of the penalty to be assessed, vary inversely with the ability to corrupt the very processes responsible for disciplining judicial behavior. The relative bargaining power of the judge and competing litigants—degree of financial need, resources, competition—all affect the level of corrupt practices. However, this approach does not focus primarily on the conditions that give rise to these incentives in the judicial process, including the opacity of the process itself, which lowers the risk of detection.

Justice institutions, courts in particular, have monopolies over some forms of decision-making. In some instances, they also may have broad discretion. For these types of cases, it becomes more important for the assessment team to pay attention to risks and incentives for corruption.

It is essential to distinguish identifying the problem from identifying its source or its solution. Insufficient pay of judges, their staff, and other justice sector employees often is mentioned as the cause for corruption. Frequently, the solution suggested is to raise pay. The level of pay may be a *proximate* cause for corruption and certainly is an issue to be assessed. However, research suggests that the *underlying* cause of corruption is primarily institutional weakness. It manifests as poor monitoring systems; political intervention in appointments and promotions; and a variety of informal rules, incentives, and cultural expectations. The presence or absence of corruption is influenced most significantly by the existence of meritocracy, quality of internal administration, and actions of external watchdog agencies—press, private bar, and civil society groups.

In Slovakia (Anderson N.d.), the courts were identified by businesses, households, and public officials as slow and largely corrupt. Enterprises identified slowness of courts and "low executability of justice" as severe obstacles to doing business. Both households and enterprises reported that they frequently encountered bribery in their experiences with courts. Moreover, there was a widespread perception that students could not gain admittance to law schools without paying bribes.

The Nigeria corruption experience index (UNODC 2006) also confirmed the strong relation between delays and corruption. Commissioned by the United Nations Organization on Drugs and Crime (UNODC), this report was the result of a survey of courts users' experience with corruption in Nigeria. The report found that, in Nigeria, delays are a compelling incentive for court users to accelerate the procedure by paying bribes. In fact, a delay often is an implicit request for a bribe in exchange for an unanticipated service.

Procedural complexity also may facilitate corruption. If cases take a long time to be dealt with in court, the passage of time may increase the opportunity for, and the likelihood of

incidents of corruption. Reportedly, in Nigeria, bribes were paid in connection with applications for bail, institution of proceedings, issuing of summons to defendants, interrogatories, delivery of judgments, and obtaining certified copies of proceedings. Persons who had to return to court several times for the same cases were the ones more frequently asked to pay bribes.

In analyzing and developing remedies for corruption, the first step is to determine the incidence of corruption and the forms it takes. Experience surveys will be useful here, although as noted they may over- or under-report certain forms. Generally, it is easier to establish the presence of petty corruption than of state capture and influence.

The second step is to gain insights into the proximate and underlying causes of each type. Well-constructed experiential surveys (asking people not only whether but also when they paid bribes) can help here. However, these surveys will have to be supplemented with other sources of data: observation, interviews, focus groups, and analyses of real cases. Case analysis rarely will provide more than inferential evidence, but the analysis can suggest the likely locations of the opportunities for corruption.

Frequently, certain types of petty corruption are easiest to attack first, because they benefit primarily low-level players. However, it is advisable to remember that, *ultimately, corruption is symptomatic of a poorly functioning system.* Consequently, reducing its most pernicious forms will require myriad individual steps. They will need to reduce delay and improve the way the system works. Particularly important is improving accountability and transparency by:

- Ensuring merit-based systems for judicial appointment, promotion, and disciplinary proceedings, as well as adequate judicial salaries and training
- Promoting transparency in proceedings and through publication of judicial decisions
- Prosecution of some high-profile corruption cases (Anderson and Gray 2006).

Because so many actors may have strong incentives for corrupt behavior, the team must be careful to identify sources of support and resistance as means of selecting entry points.

The research quoted strongly suggests correlations between the three most common complaints. Delay impedes access; it also creates opportunities for corruption. Corruption itself has the effect of reducing access. Attacking one complaint will most probably affect the other two. These correlations are relevant when the team formulates recommendations. Chapter 6 discusses how to formulate recommendations in detail.

6. REPORTING AND DISSEMINATION

Writing and disseminating the assessment report mark the culmination of the assessment process. While writing the report may not begin until this point, constructing the report or thinking about how it will be written often begins during the research phase, as findings are evaluated and tested. In this sense, the work at the end flows directly from the data collection and analysis phase described in Chapter 3. Indeed, the types of analysis described in Chapter 3–background, problem identification, findings–will constitute the bulk of the report's content. The main task at the last stage is for the team to assemble all of its findings and analysis and distill them in written form. The goal is to narrow down the data collected–choosing what to leave in and out–to present the most relevant findings in the most effective way. This exercise is one-third organizational, one-third analytical, and one-third political.

The aim of Chapter 6 is to provide guidance on each of those fronts. The *organizational* side is relatively straightforward and entails organizing the content to present the findings in the most effective way. The team will need to pay attention to which content appears in which part of the report (background, findings, appendix) and the format and style in which the material is presented (description, text box, figure, table).

The main *analytical* challenge is to draw *clear conclusions* that go beyond overly general findings and point to *workable solutions* to the identified problems. However, the results might exhibit great ambiguity. This *ambiguity itself might be the main finding and conclusion.* If so, the challenge is to provide solutions and clear paths that go beyond the ambiguity. A solution might be to simply flag the uncertain areas so that reformers are attuned to them in follow-up activities. It also might entail recommending a focused follow-up research plan–but being sensitive to a frequent perception that the goal of people performing these assessments is to generate more assessments! The *political* side might be the thorniest with no clear solution. In short, the report will have multiple audiences, and it is usually not possible to satisfy everyone. The team's task then becomes to head off potential criticisms and dissent. There will be a tension between trying to satisfy as many constituencies as possible but not compromise the integrity of the report's findings.

One factor to be considered when producing the report is that *conclusions are time sensitive.* The assessment may have begun to influence the existing state of affairs before the final report is produced. The longer the time for the report to be disseminated, the more likely the report is to lose its relevance. The risk is that by the time the report is produced, events will have passed it by and its content will no longer present an accurate account of prevailing conditions and their solutions. Thus, the team needs to get the assessment out quickly but without compromising the quality of the analysis.

The following testimonial from an assessment team member vividly captures all of the challenges of the report stage:

<div style="border:1px solid black; background-color:#d3d3d3; padding:10px;">

Box 6: Reporting Challenges

"Speaking from experience there are two problems here. One is the CMU's nervousness about criticizing certain agencies and actors (not necessarily the judges). I was forced to eliminate comments which suggested the national president was leaning on the judges and that the external disciplinary board had threatened some of them. The local rep also wanted me to be more critical of the Attorney-General who he thought was a crook. Also getting permission to release the document can be iffy—our rules say that if the counterpart agency (which may not be the court) does not respond within a reasonable time we can release. However the WB doesn't recognize positive silence in fact."

</div>

Audience

The report must, first of all, meet the needs of its readers and answer the questions in their minds. The report has multiple audiences: local counterparts, local stakeholders, Bank staff, experts, decision-makers, and system users. Moreover, for part of its audience, the report will not be in their native language. The report will have to be written with all of those audiences in mind. Be aware of the different needs that each may have. Some readers may have in-depth knowledge of the subject; others may be decision-makers without specialized technical knowledge. The report should explain why something is a good idea, use clear, straightforward sentences and not make assumptions about the readers' understanding. It should use generic terms for legal concepts and explain to which local institutions they refer. If the report is clear and interesting in both content and presentation, the team need not fear losing any of its readers, whatever their level of understanding.

World Bank reports should follow the "World Bank's Author's Guide" and "Style Guide" for writing and preparing World Bank manuscripts.[54] The latter provides guidance on approved word usage and formats to use in preparing manuscripts for publication. The "Style Guide" draws on *The Chicago Manual of Style*, 15th edition, with modifications to reflect the special nature of the Bank's work.

Consider whether translations can be done, taking into account that they take more time, involve considerable cross-checking with the original (a problem if the initial drafters do not understand the native language), and, of course, add to costs. For some audiences, presenting the report with an explanatory note that is not part of the report may be a solution.

Basic Elements of the Report Content

The tips to follow for organizing the content of the assessment report are fairly basic. However, the importance of a well-structured report cannot be stressed enough. A good

---

[54] These can be downloaded at
http://intranet.worldbank.org/WBSITE/INTRANET/UNITS/EXTAFF/EXTOP/0,,contentMDK:20820785~menuPK:2902133~pagePK:85120~piPK:445555~theSitePK:84816,00.html

report will have a clear logical structure with visual signposting and section headings to show where the ideas are leading. This section provides a suggested outline to follow in structuring the report's content and can be used as a checklist. Guidance is provided specific to each of the reports sections, which are:

- Executive Summary
- Introduction
- Background Information
- Methodology and Limitations of Analysis
- Findings
- Recommendations
- Appendixes
- References.

Executive Summary
This section of the report is the one that the team writes last but that people will read first. It also often is all that they will read. Thus, the Executive Summary should be brief. It should contain a summary of the other sections and as many of the core findings and recommendations as possible. It may rarely be possible to sufficiently support conclusions in the summary. When it appears necessary, footnotes can be added that direct readers to the relevant sections of the report.

Introduction
The introduction should provide a brief description of such things as the assessment's topic, how it was chosen, and the initial impetus for performing the assessment. Directly related, any assumptions made about the role of justice systems and their relation to development and the reform process also should be spelled out clearly. Is the assessment proceeding on the basis that well functioning justice systems are essential for growth and efficient markets? Is it proceeding on the basis that access to justice is a fundamental right and essential in reducing poverty? Is a well-functioning justice system considered an essential public good in its own right? There also should be a basic narrative about the scope of the assessment, for instance, who performed it, over how many days, and which regions and institutions were covered. A more detailed list of team members, interviewees, and sites visited can be included as an appendix.

Background Information
The material in this section should be selective. It should provide a description of the legal and judicial system, and of any other relevant country conditions (historical, social, cultural, or political). To save limited energy and time, the description should focus only on the background information needed to support the subsequent analysis. Many written assessments err in including too much, turning what should be a few paragraphs or pages into a lengthy treatise. The counterparts already know most of this material, and its extensive repetition is one of the reasons for diagnostic fatigue. Their common complaint is, "We already have too many studies, and none of them tells us what we don't already

know." A lengthy introduction can present its own problems, provoking dissent about details that are in the end irrelevant to the assessment itself.[55]

The obvious solution is to put some of the information that, while relevant, is less immediately essential to the report's analysis into an appendix or background papers, and leave only the most critical elements in the main report.[56] It may be most useful to present this material in the form of tables or figures rather than as text. For example, a chart showing the organization of the state system can replace pages of written description. It may be possible to do the same for nonstate bodies. Lengthy paraphrasing of constitutional and legal provisions should be avoided altogether, both here and in the findings section.

Description of Methodology and Limitations of Analysis
This section of the report provides the opportunity to describe in more detail the approach taken in performing the assessment research. Demonstrating the methodological soundness or self-identified limits could go a long way to ensure the success and perceived legitimacy of the assessment. Demonstrating credibility will be especially important for the points that will be more susceptible to criticism because they expose previously unknown or controversial issues.

At a minimum, the section should provide a brief overview of the research methods used. This overview is likely to include informant interviews (and types of informants), observation (of what and when), regions covered and why they were chosen, analysis of available statistics, and review of basic documents. To buttress the descriptions provided in this section, the questionnaires used for interviews and surveys and the databases and documents consulted can be included in an appendix.

The section also should discuss what the assessment was unable to do and how this inability might affect the results. This discussion would highlight issues that could not be probed more deeply, types of informants who could not be consulted, and limitations in the methodology as they arose. A common limitation could be lack of time and resources to research more deeply.

To be clear, however, the limitations of the research need not necessarily undermine its results. The point of underlining possible limitations is to ensure that the assessment results are neither perceived to claim more than they can, nor interpreted to claim more than they do. For instance, an assessment may be able to cover only urban centers and not rural areas. Highlighting this fact does not take away from findings about the urban

---

[55] Several years ago, another donor encountered just such a problem when the chief justice took issue with some of the historical interpretations and refused to agree to the report's publication.
[56] This is commonly done with other types of ESW, for which background material necessary for other readers is largely a waste of time for, and may provoke unnecessary conflicts with, the counterparts. The conflicts do not arise because the material is sensitive but simply because the counterparts may not agree with a description aimed at outsiders and framed in terminology that the latter will understand.

centers. It merely underscores that the findings do not necessarily reflect realities in the rural areas, which may or may not need to be the focus of their own follow-up research.[57]

Findings
The findings section is the most important part of the report. Although it may not be the most widely read part, *the findings section contains the main presentation and analysis of the data. These serve as the basis for any conclusions and recommendations.* For this reason, the legitimacy of the report may rest on the soundness of this section. As a general rule, findings should be presented in descriptive terms and visibly backed by the research. They should not be judgmental in the sense that any judgment about the quality or shortcomings of a specific state of affairs should be based on a reference to a pre-identified set of criteria, with some explanation for why these criteria were chosen in lieu of others.

The content of this section can largely mirror the order and types of analysis discussed in Chapter 3. The *analysis* includes the following subsections:

- *Discussion of preliminary problem.* What were the initial findings based on the preliminary problem? In what direction did it lead the analysis? Were preliminary causes of the chosen problem identified by this subsection, and were they probed more deeply in the following stages of the assessment? The discussion also should note any initial impressions or hypotheses about causes that were *not* validated by the later research.

- *Summary overview and analysis of operations of state of affairs.* This subsection will provide a picture of the current state of affairs based on the results of the assessment. It will answer the question "What is out there?" As noted in Chapter 3, it may be helpful to organize the data according to the experiences and information relating to individual organizations, categories of actors, or regions, prioritizing the most relevant ones. A more generalized picture and trends then can be distilled from the individual cases.

  Alternatively, it is possible to present the results according to the generalized findings or trends, but present only the most salient individual cases. Most salient would be the cases that are most indicative of the trends witnessed, or that are exceptions to the trends; and the implications of these exceptions for the more general findings. Whether to proceed from more general threads or a case-by-case analysis will be determined largely by what the assessment's results most readily lend themselves to.

  Regardless of the method chosen to present the results, both types of analysis–case by case and general threads–will have to be performed. The decision on how to present them in written form then becomes largely about style and expediency of writing. This subsection is also a prime spot for many of the visual techniques of presenting

---

[57] Whether the research should have restricted itself to urban centers and why it did so also may be issues that need to be addressed.

data discussed below (maps of the sector or geographic distribution, tables or figures of budget allocation).

- *Analysis of problems posed by the state of affairs.* The description of current conditions should lead to a discussion of the problems posed by this state of affairs. The discussion on findings should clearly differentiate the first type of analysis (how things are) from the second type, which involves identifying problems (unsatisfactory outcomes or impacts) by applying a predetermined standard or benchmark. In this sense, when discussing the general findings on problems, the report should be very explicit about which standards are being applied, and how or why they were chosen.

  The general point to bear in mind–which will flow throughout the findings section of the report–is that the data collected might not yield a definitive conclusion about what problems are created by the state of affairs. It is likely to take far more time and resources than would be available to determine whether, for example, delays, bias, and negative impacts on the poor or the business community are as extensive as the interviews and a few anecdotal examples might suggest. Ambiguous or uncertain results should be confronted head on and presented for what they are. This in turn likely would mean being more cautious or tentative in making recommendations and devising strategies as a first step in probing the ambiguities more deeply at a later stage.

- *Conclusions and analysis of proximate and underlying causes.* After describing the state of affairs and problems arising from it, the next step is to discuss the proximate and underlying causes of these problems. This exploration will serve as a basis for the recommendations in the next section of the report (solutions should be shaped by and address these causes). The analysis should follow the methodology described in Chapter 3, section 5, on how to isolate and prioritize causes from the data. This subsection will describe this process of analysis and its outcome. Again, the goal is not necessarily to come to a thoroughly conclusive set of causes, but to suggest how clearly a certain set of findings imply one or another set of causes, proximate or underlying. Again, the proper conclusion to highlight might be that no clear cause can be concluded based on the data.

- *Comparative evaluation.* The findings should not be discussed in a vacuum. To add more context to the discussion, the findings should be compared and contrasted to other experiences. Two types of experiences stand out: (1) What problems do the issues discussed present for users and for society as a whole? (2) How do the problems compare to those of similarly situated countries? In the first case, relating the topic to the experiences of the society as a whole puts in perspective how this issue relates to other pressing concerns confronted daily by the population. There may be a pressing need for reforms in the justice sector, but how does it compare with needs in other essential sectors such as health and education? The answer is by no means obvious but will relate directly to the type and scope of recommendations made. For example, wholesale calls for reform may ring hollow if other crucial sectors are facing equally deep and widespread problems. In addition, if an issue

affects a small sector of the population but in a serious way, how should recommendations be formulated when the majority of the population may be preoccupied with other issues?

In the second case, comparative experiences from other countries will help to give a sense of proportion to the issue. Similar experiences with the same problems are likely found in similarly situated countries. From a practical standpoint, relating a problem identified by the assessment to outside experiences could help to defuse a sense that the report is being overly or unfairly critical. How the same problems have been confronted in other contexts also can be a source of inspiration for solutions. Finally, looking to other experiences can help to isolate what may be idiosyncratic to the assessment's particular context, which in turn can help to ensure more context-sensitive and tailored conclusions and recommendations.

Recommendations
The goal in making recommendations is to provide counterparts with possible solutions to address the problems identified and analyzed by the assessment. Great weight may be placed on this section by many readers as their interest in the report may extend only to discovering these solutions. Similar to many medical patients, reformers may not necessarily be as concerned with the diagnosis or causes of a problem as they are with the remedy the doctor prescribes to cure their ills.[58]

Nonetheless, formulating useful recommendations is an art. The biggest challenge is to devise solutions or next steps that provide a sufficiently focused and concrete strategic basis to guide reformers. Common complaints about recommendations are that they do not tell the audience anything new and, in this sense, are too general or obvious to translate into action or next steps; that they reflect the author's own preferences or biases without a clear link to the analysis of problems; or that they are an unrealistic laundry list of solutions that lack any indication of how to prioritize them or of how they relate to one another.

Formulating useful recommendations will ideally flow directly from the analytical process.[59] The proposed reform strategy should be embedded in an analysis of the principal causes of poor performance and the recommended means for dealing with them. A realistic and well-prioritized list of recommendations would distinguish between the levels of importance and urgency of problems that need to be addressed; the timeframe for activities and achieving results (short-, medium- or long-term); and any possible "quick fixes" or easy victories for reform (or, to use a current buzzword, "low-hanging fruit"). An activity that (1) tackles important or urgent issues, (2) can be addressed with a

---

[58] In addition, as with many patients, they may not necessarily follow the prescription fully, such as not taking the full dose of antibiotics; or may focus too much on the quick fix elements, such as taking the antibiotics, while neglecting the elements that are harder to follow and have a less obvious impact, such as getting enough rest and fluids.
[59] Recommendations cropping up during the analytical process should be parked for future use. The parked recommendations may then find their way into the report.

short-term activity and (3) needs few resources generally is considered a low-hanging fruit to be reaped quickly and easily.

The recommendations also should highlight the possible risks and obstacles in the path of those reforms. For more detailed guidance on formulating recommendations, please see Chapter 3, section 6.

Finally, recommendations should be tailored to what existing legislation and institutional structures realistically permit. Too often, assessments recommend new legislation as a reform solution, but this poses two difficulties. First, legislation requires collective, often lengthy, political action, and the political process often delays or dilutes the substance of the proposals. Second, the call for new legislation often implies that the problems lie mainly in the rules, not in the conditions of institutional arrangements and incentives for primary actors. Past assessments have found that the appropriate laws were in place but they were not implemented. Alternatively, if implemented, they were not followed. Quite often, the problems have other causes.

Appendixes
Appendixes are used to capture any residual information that, while potentially interesting to readers, may not be entirely relevant to the main findings of the report or may detract from its narrative flow. Appendixes flesh out content appearing in the main body of the report. Specifically, they can be used to compensate for the different knowledge or expertise levels of the report's audiences. For instance, a detailed description of the justice system will be less relevant to local actors than to a foreign audience.

Appendixes also reinforce the rigor and comprehensiveness of the report's analysis. For example, including questionnaires from interviews would be misplaced in the body of the report but entirely appropriate in an appendix. The appendixes frequently total more pages than the main report. This fact invites the perception that they are a dumping ground for leftover material. While there definitely is a broader scope of discretion as to what to include in the appendixes, authors should carefully exclude tangential (albeit interesting) information and include only information that enhances the messages in the main report.

Box 7: Appendixes that should be included in a Justice Sector Assessment report

- People interviewed and their affiliations
- Regions of the country visited
- Other materials consulted (databases).
- Assessment tools (informant or focus group discussion questionnaires).
- Explanation of methodologies: descriptions of how samples are drawn for surveys or case file analysis, criteria used for evaluating judgments.
- Additional background as relevant
- References

Formatting Report Content

In addition to thinking about what to include in the body of the report, the authors also need to give thought to *how to present* the report's content. *The goal is to keep the report's content on message.* Visual, narrative, and other techniques can enhance the report's comprehensibility by making its content as easily decipherable and accessible to the reader as possible. In other words, not all of the content should be put into written form. These techniques should facilitate the job of the reader who, realistically, might give the report only a cursory review. In this sense, the authors need to be judicious in choosing which technique to use, where, and how often. As can be expected, there will be a balance between when a technique enhances or detracts from the report's content. For instance, too many large blocks of text could cause the reader's eyes to glaze over, whereas too many boxes or charts could make the report too busy and hard to follow, without much of a narrative thread to hold it together. In addition to their visual impact, the different techniques will strengthen the report's content by providing an additional perspective on a given issue.[60]

The following is a list of suggested techniques to use in presenting the content of the report. Examples of each of these techniques are found throughout the body of this handbook. In preparing the handbook, the authors attempted, as much as possible, to follow the advice on content provided herein.

Diagrams, charts, graphs (technically known as *figures* and *tables)* are all useful for avoiding lengthy descriptions, particularly when the audience can be expected to be familiar with the information represented in the graphic. Graphs and charts (bar graphs, pie charts) are particularly good for giving a general overview and visual representation of statistical results. Written description can then serve to highlight a few of the salient findings. Organizational charts and flowcharts, as noted above, are particularly handy for explaining things like institutional or organizational structures or steps in a process (the life of a legal dispute), respectively.

---

[60] For instance, as between a graphic representation of a state of affairs and an anecdote describing a particular instance of that same reality. The authors should seek the optimal balance of where and how often to use a technique.

Text boxes can highlight particularly salient information or information that is directly related to the surrounding text but that does not flow comfortably with the narrative. The added benefit is that text boxes can break up long stretches of text. Text boxes are good visual vehicles for case studies and anecdotes.

Cases studies and anecdotes both serve as illustrations and enrich the discussion of an issue by going into more detail about a particular set of experiences. The case or anecdote should give a deeper dimension to the discussion, by either bolstering or serving as a counter-example to a more general point. For instance, the discussion might usefully describe the experiences with a particular courthouse, or in a particular region, and how actors there coped with a particular issue, either to success or endless frustration.

The obvious risk with using case studies and anecdotes is that a particular instance can get mistaken for a general rule. In other words, the authors need to take care not to overstate the significance of an individual case. While the case study or anecdote may provide an accurate instance of a general trend or finding, it also could represent what is in reality an idiosyncratic state of affairs. Alternatively, the case study or anecdote might appear to affirm a rule but, in reality, describe a "false positive."[61] In this case, the illustration would misattribute a general rule to the specific case. The general tip here, which may seem obvious, is to ensure that the underlying analysis is sound, and then pick an example or case that flows clearly from this analysis.

The difference between an anecdote and case studies is that *case studies investigate a generalizable issue whereas an anecdote exemplifies a specific point.* The case study's description is more impersonal and systematic. While its findings should be extrapolated with great care, they usually are supported by other types of evidence, such as aggregate statistics or surveys. In contrast, anecdotes generally are more informal and illustrate rather than "prove." The main impact of a well-placed anecdote is to bring a more direct or human dimension to the issues being discussed. It gives the reader an example and added feel of how the issues affect the individuals confronted with them, such as a court user trying to navigate the system or a court administrator trying to cope with a heavy backlog. The anecdote could be a striking quotation from an interview or a longer paragraph or story. The key is to employ anecdotes that bring out an added dimension of an issue, which might not be immediately obvious from the other descriptions; or that capture a particular point with the succinctness expressible only through the personal experiences they describe. If not skillfully selected, anecdotes can come across as trite or heavy handed.

Footnotes: Academic writers, as well as others, may be inclined to use footnotes to (1) attribute direct quotations, statistics, and general information drawn from other sources; (2) provide supporting documentation for statements that might be questioned by readers;

---

[61] As an example of a "false positive," consider a court with a low backlog. The conclusion could be that this court "works well," in contradistinction to courts that have chronic backlogs. Alternatively, the court could suffer from the same problems as the other courts but has a reduced backlog because court users have given up using the courts entirely. Far from affirming the opposite, this latter example would represent an even "worse" state of affairs than in the other courts.

(3) cite related material consulted by the team or of possible interest to readers; and (4) define terms, elaborate on points, or give examples in a way that does not interfere with the flow of the text. Most of the footnotes used in this guide are of the fourth type.

Footnotes are useful but should be used wisely. Their format should follow the social science rather than the legal style:

1. The author-date system (author's last name and year of publication embedded in the text) should be used as much as possible.
   a. Example: "…other authors (Smith 1999; Jones, 2006) have written on this point…"
   The complete citation must appear in the References.
2. References within footnotes should follow the same abbreviated format as 1. above (author's last name, year, and, if relevant, page number). The complete citation must appear in the References.
3. If the author-date system is being used to cite references to media, notes formatted as footnotes or endnotes should explain topics other than author and year of publication. Either footnote or endnote format is acceptable, but the same format should be used throughout one document. These notes should be only a few lines long. If the note mentions any form of media, the complete citation must appear in the References.
4. For the remainder, including citation of laws, interviews, and internet material, additional guidance can be found in the "World Bank Editorial Style Guide"[62] and in *The Chicago Manual of Style* (15th ed.).
5. If the handbook has no References section, then every reference to a publication in the text must have either a footnote or endnote that contains the complete citation.

Writers from whatever discipline are advised to curb their usual enthusiasm for frequent and lengthy notes (whether embedded, endnotes, or footnotes). Although adequate documentation is desirable, authors should keep in mind that this handbook is not a paper for an academic journal and that excessive referencing will not appeal to many readers.

Ensuring Report Receptivity
A concern underlying the whole report writing exercise is to ensure support for the report and its recommendations among its intended audiences. The above tasks, of deciding what content to include and how to present it, will play a role in ensuring that the report is well received in that they can make the content as accessible as possible to the reader.

In addition, there is a political side to this exercise, which has a much less certain methodology and outcome. The challenge is to be critical and constructive at the same time. There is likely to be some apprehension about the report within the reform community born of a common resistance to critical views of outsiders. While the report might be intended, first, for an internal World Bank audience, it also needs to be sensitive

---

[62]http://intranet.worldbank.org/WBSITE/INTRANET/UNITS/EXTAFF/EXTOP/0,,contentMDK:20823865~menuPK:2902133~pagePK:85120~piPK:445555~theSitePK:84816,00.html

to local audiences. Again, anticipating the concerns of different constituencies without watering down the report's content can be a fine line to walk. However, care is called for here because the final impact will hinge, to a large extent, on the enthusiasm and willingness of the local reform constituency to act on the report and advance its recommendations.

Ensuring receptivity to the report is not an exact science. Some basic tips are described below. This list is not intended to be exhaustive. Deciding which tips to employ will be a judgment call based on the specific context confronted. Ensuring receptivity to the report blends into the subsequent topic of devising dissemination strategies for the report.

1. *Share drafts* with key stakeholders, prior to finalizing the report. As with any draft, soliciting feedback can improve its content. For instance, there might be slight inaccuracies in how an issue is portrayed. Reading the report might spur reflection and lead to additional insights and ideas from stakeholders on issues missed during a first round of interviews.

   Sharing the draft also is a concrete method to demonstrate that the stakeholders' inputs are valued (provided input is not done perfunctorily). Stakeholders—as both reformers and actors in the system—often express the wish to see the report following an interview. Sharing drafts also can help which gauge findings and recommendations from the report will be controversial and at which points different stakeholders will have an interest in the report's findings.

   Ideally, sharing drafts should not affect the content of the findings and recommendations. Unfortunately, there have been cases in which various stakeholders have tried to thwart completely the release of a report due to its contentious findings, for example, when they are deriving large benefits from the current problematic conditions. How to navigate these cases will require tough judgment calls based on the contexts, with possibly no easy solutions. This is all to underscore the potential impact of an assessment report and the critical need for sound analysis in the assessment.

   Sharing drafts also can be used as part of the dissemination strategy for the report. For instance, a workshop can be held with stakeholders to provide an overview of the findings and solicit input to feed into the finalization of the report.

2. *Use coauthorship with local counterparts* to help ensure greater ownership of the final product. Coauthorship here would mean involving counterparts in various parts of the assessment, especially at the data collection phase, and presenting the report as a joint product. That is not to say that the actual writing of the report needs to be shared by everyone. Too many writers, whether with local counterparts or otherwise, can lead to a disjointed report, with a combination of inconsistent narrative styles and overlapping discussions. A decision will have to be made on who handles which part of the writing and under whose direction.

3. *Use constructive and diplomatic language*. Findings should be presented in descriptive terms and clearly backed by the research. Assessments should not be judgmental. Often, ensuring positively oriented assessments entails a subtle shift in vocabulary: for instance, replacing "crisis" with "challenge." In other instances, the choice of language may depend on whether the constructive effect of diplomacy outweighs the value of candor in assessing findings, or vice versa.

4. *Discuss comparative experiences* and shared challenges in other countries, especially in the same region. These discussions will help to put the problem into perspective by showing that the problem is neither unique to, nor necessarily more severe in, any given country.

Report Publication and Use of Its Contents
Generally, the World Bank's rules and contracts signed with consultants make the report and the research on which it is based the property of the World Bank. Consultants often use their own research, sometimes without permission, but contracts also can be amended to enable them to do this, possibly with prior Bank review.

However, Bank rules also favor report publication based on prior consultations with the primary counterparts. Usually, the official counterpart, even for a justice assessment, is the Ministry of Finance or Planning, but obviously the team would wish to consult with the judiciary and any other institution covered extensively. The issue can be delicate because Bank rules do not specifically require counterpart acquiescence—only consultation and consideration (not necessarily incorporation) of suggestions.

On the other hand, few CMUs would publish a report, even in amended form, to which a counterpart objects, or even—a remarkably frequent situation—when it offers no comments, positive or negative. Positive administrative silence is not practiced by the Bank in these situations. Thus, a final chore for the assessment team, and especially its Bank members, may be to consult with the official and other counterparts to ensure that they respond to the request for comments and stipulate their conditions for wider dissemination. Once the report is published, its contents can be used and cited in the same way as any other copyrighted document.

APPENDIX 1. SUGGESTED OUTLINE FOR ASSESSMENT DATA COLLECTION AND ANALYSIS

(More information on types of data and collection methodologies is given in chapter 4.)

I. Background Material

In this preparatory stage, the following types of information can be collected from published and internet material; agency reports; and personal, telephone, and electronic interviews with country experts.

A. Political and socioeconomic characteristics of the country that encompass:

1. Population and population density; major ethnic, religious, and other divisions; location; geographic dimensions and main characteristics; size and composition of economy; brief history
2. Political characteristics: Current regime type, brief history of how it got there, any ongoing political strife, other factors affecting stability
3. Characterization of developmental situation: Per capita income; economic trends, sources of growth; level and distribution of poverty (regional? ethnic? rural and/or urban?); basic economic, social, and political challenges faced.

B. Preliminary overview of justice system to identify:

1. State system: Major legal tradition(s), general structure, and component organizations
2. Other "modern" auxiliary or informal organizations: Mediation centers, bar organization, nongovernmental organization and university service providers, registries, private bailiffs, notaries
3. Traditional nonstate system: Less formally recognized religious courts and community justice
4. Inventory of basic legal framework.

C. Identification of high-profile conflicts and the extent to which they appear addressed by the sector agencies:

1. Crime
2. Land disputes
3. Ethnic cleavages.

II. Preliminary Problem Identification

Problem identification usually takes place at the beginning of the in-country field work. However, relevant information also may be gathered prior to the mission, especially from experts residing outside the country or from written sources that the team has identified in the preparatory work.

A. If the problem is preassigned, then a short exploration of its nature, drawing on documentary sources and informant interviews, preferably from a variety of sources, but not yet organized as formal surveys or focus groups

B. If the search is more open ended, use of available documents and surveys and a series of unstructured interviews with a variety of sources to begin to define and rank perceived problems

C. Some initial testing of hypotheses derived from either A or B against evidence collected in part I and initial observations to determine whether they are on track or whether additional work is needed.

III. More In-Depth Review of Justice Sector Institutions

The selection of the institutions and the extent of attention to each are guided by the problem definition(s). Sources of information include documents (including laws and published, internet, and unpublished reports); informant interviews, observation, statistics (official databases and, if necessary, statistics generated by the team), focus groups, and less formal participatory exercises.

A. Formal institutions:

- Courts or court systems, administrative tribunals, prosecution, defense, police, state-managed mediation services, other
- Details of organization (a chart may be most useful) and of overall powers and duties, as well as internal distribution of labor
- Body(ies) responsible for organizational governance and administration: Composition, powers, focus of operations (day-to-day administration, policy setting, planning?)
- Human resources: Overall number; major job categories and distribution of work force among them; employment conditions (salaries, tenure, career system or not); means and conditions for selection; performance monitoring, if any; skill levels and training programs
- Geographic distribution of work units, employees, and workloads
- Budgets, sources, how set, functional and geographic distribution
- Other resources—infrastructure, ICT equipment, vehicles—and their distribution
- Law/rules ("normative framework") governing operations, process required to change them, and a brief summary of their known or likely impact on real operations
- Rules for accessing their services and likely impact of the above on access; potential barriers such as geographic distribution, payment of fees, need for legal representation

B. Nonstate (for example, private or NGO-run mediation services; private security services and bailiffs) and traditional institutions involved in conflict resolution:

- When relevant, many of the categories under III.A and information sources can be used.
- Additionally, information will be needed on nonstate entities' legal and real interface with the state system—are they recognized? If so, is their jurisdiction further defined? How are conflicts of jurisdiction handled? Who is entitled to use them and who actually uses them?
- What types of conflicts usually are handled; types of rulings made; if possible, some estimate of workload and growth or decline in recent years.

C. Related state and nonstate institutions affecting operations of A and B, the private bar and any bar associations, government legal services (those representing the government in litigation and providing legal advice), notaries, registries, credit bureaus. Selection and extent of focus will hinge on the problem definition(s).

- A truncated version of the information collected for III.A can be used. Basic data on organization, governance, size, and workloads always will be relevant for any entities covered. Additional data will depend in large part on the types of problems being explored.
- Additional information for the institutions covered will focus on the type of inputs provided to the conflict resolution process, its adequacy, and its impact on the quality of outcomes.

IV. Review of System Operations

Again, the problem definition(s) will guide the focus. If the issue is delay, the team will attempt to verify and measure its existence and dimensions and begin to track possible explanations for the impressions or reality of its existence. If the initial identification has identified two or more problems, the focus will broaden, but the problems still will provide some guidance as to emphasis.

To review system operations, somewhat different sources of information are used. While interviews, documents, and official statistics may help, surveys, focus groups, observation, legal analysis, and reviews of actual cases will be more important. The sophistication of the methodology will depend on the resources available.

A. Workloads and productivity

Initial questions are what exactly the system and its organizations do and how the work is distributed. Where available, organizational databases can be a primary source of this information, supplemented by documentary sources, observation, and interviews. In the absence of statistical data, the latter three sources, or work unit records and registries, may have to suffice. Among the issues that can be explored at this global level are:

1. Match between geographic distribution of work units and demand

2. Composition of demand (and if time series data exist, changes over time) globally and by geographic area
3. Gaps between supply of services and demand, globally and by functional area.
4. Signs of problems originating in insufficient inputs from key actors and organizations; for example, significant differences among number of crimes reported, investigations completed, and cases taken to trial; or between amounts awarded in judgments and payments

B. Case trajectories and outcomes

It always will be useful to track actual case processing ("cases" here to be understood as the organizational product—which for police might be an investigation; for defense, a case defended) by using organizational records, interviews, and observation. As this is time-consuming work, selection of case types should be based on findings from part IV, section A above and areas identified in the problem analysis. Whether working with a few cases or some sort of sample, four steps are recommended:

1. Development of a flowchart of normal steps in the cases studied. Flowchart may include a comparison of legally defined steps and those actually occurring.
2. Identification of alternative routes and outcomes, and the reasons for each.
3. Qualitative evaluation and, if possible, some quantification of the likelihood of different routes and outcomes.
4. Identification of areas of poor performance, obstacles, and bottlenecks; and of desirable changes in the patterns identified.

C. Exploration of problems identified

Where the above steps do not provide sufficient information, the following methods can be added to expand the explanations for an illustrative set of typical problems (see also chapter 5).

1. Delay: Use of aggregate statistics, case file analyses, surveys, observation, informant interviews, and focus groups
2. Quality of decisions: Observation, informant interviews, analysis of legal provisions, legal analysis, review of a sample of actual decisions
3. Enforcement: Court records, informant interviews and focus groups, legal and procedural analysis
4. Corruption: Surveys, informant interviews, observation. While actual corruption may not be observed, vulnerable points can be identified. As a consultant once said, "The man who sits by the door gets the first chance at the bribes."
5. Access: Informant interviews, surveys, aggregate statistics (to draw inferences from geographic origin and types of issues covered), legal and procedural analysis (to identify such obstacles as filing and attorney fees, rules that may

add other costs or work against certain parties, or informal practices that may have similar effects), case analysis to determine who usually wins.

V. Comprehensive Analysis of Partial Findings, Prioritization of Problems and Their Causes, and Identification of Areas for Interventions

In this stage, additional data will be collected only to fill gaps. Instead, the entire team must compile its findings to determine (1) whether the problems initially selected are the correct ones; and if not, which should be substituted; (2) what the principal causes are; and (3) what types of measures might be introduced to improve the situation.

Among the factors to be considered in identifying causes and devising remedies, the following usually are most important:

A. Legal framework: May complicate or obstruct conflict resolution, impede access, or alter actors' incentives.
B. Organizational weaknesses: Inadequate resources; poorly prepared, motivated, or under-supervised staff; structural factors that facilitate external pressures and influences.
C. Failures of interorganizational coordination or inadequate input from one or more actors.
D. Societal or cultural practices or conventions not based in the law that may nonetheless affect sector operations. These commonly include biases against certain groups, but they also extend to certain unwritten standards (for example, "protection" of women or of one's honor, the responsibility to honor kinship ties).
E. Incentives: The factors, usually drawn from one of the above sources, that determine how actors actually perform their officially defined duties.

VI. Recommendations

Recommendations will be based on a prioritization of problems and on a series of short-, medium-, and long-term corrective measures.

Universal Declaration of Human Rights

10 December 1948, General Assembly Resolution no. 217A(III), U.N. Doc. A/3

Article 10

> Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his rights and obligations and of any criminal charge against him.

Article 11

> (1) Everyone charged with a penal offence has the right to be presumed innocent until proved guilty according to law in a public trial at which he [or she] has had all the guarantees necessary for his defence.

> (2) No one shall be held guilty of any penal offence on account of any act or omission which did not constitute a penal offence, under national or international law, at the time when it was committed. Nor shall a heavier penalty be imposed than the one that was applicable at the time the penal offence was committed.

Conventions
International Covenant on Civil and Political Rights (ICCPR)
16 December 1966, General Assembly Resolution no. 2200A (XXI), U.N. Doc. A/6316

Article 14

1. All persons shall be equal before the courts and tribunals. In the determination of any criminal charge against him, or of his rights and obligations in a suit at law, everyone shall be entitled to a fair and public hearing by a competent, independent and impartial tribunal established by law. The press and the public may be excluded from all or part of a trial for reasons of morals, public order (*ordre public*) or national security in a democratic society, or when the interest of the private lives of the parties so requires, or to the extent strictly necessary in the opinion of the court in special circumstances where publicity would prejudice the interests of justice; but any judgment rendered in a criminal case or in a suit at law shall be made public except where the interest of juvenile persons otherwise requires or the proceedings concern matrimonial disputes or the guardianship of children.

2. Everyone charged with a criminal offence shall have the right to be presumed innocent until proved guilty according to law.

3. In the determination of any criminal charge against him, everyone shall be entitled to the following minimum guarantees, in full equality:

> (a) To be informed promptly and in detail in a language which he understands of the nature and cause of the charge against him;

> (b) To have adequate time and facilities for the preparation of his defence and to communicate with counsel of his own choosing;

> (c) To be tried without undue delay;

> (d) To be tried in his presence, and to defend himself in person or through legal assistance of his own choosing; to be informed, if he does not have legal assistance, of this right; and to have legal assistance assigned to him, in any case where the interests of justice so require, and without payment by him in any such case if he does not have sufficient means to pay for it;

> (e) To examine, or have examined, the witnesses against him and to obtain the attendance and examination of witnesses on his behalf under the same conditions as witnesses against him;

> (f) To have the free assistance of an interpreter if he cannot understand or speak the language used in court;

> (g) [To not] be compelled to testify against himself or to confess guilt.

4. In the case of juvenile persons, the procedure shall be such as will take account of their age and the desirability of promoting their rehabilitation.

5. Everyone convicted of a crime shall have the right to his conviction and sentence being reviewed by a higher tribunal according to law.

6. When a person has by a final decision been convicted of a criminal offence and when subsequently his conviction has been reversed or he has been pardoned on the ground that a new or newly discovered fact shows conclusively that there has been a miscarriage of justice, the person who has suffered punishment as a result of such conviction shall be compensated according to law, unless it is proved that the non-disclosure of the unknown fact in time is wholly or partly attributable to him.

7. No one shall be liable to be tried or punished again for an offence for which he has already been finally convicted or acquitted in accordance with the law and penal procedure of each country.

Convention Against Torture (CAT)
10 December 1984, General Assembly Resolution no. 39/46 U.N. Doc. A/39/51

Article 1

1. Any Torture means any act by which severe pain or suffering, whether physical or mental, is intentionally inflicted on a person for such purposes as obtaining from him or a third person information or a confession, punishing him for an act he or a third person has committed or is suspected of having committed, or intimidating or coercing him or a third person, or for any reason based on discrimination of any kind, when such pain or suffering is inflicted by or at the instigation of or with the consent or acquiescence of a public official or other person acting in an official capacity. It does not include pain or suffering arising only from, inherent in or incidental to lawful sanctions.

2. This article is without prejudice to any international instrument or national legislation which does or may contain provisions of wider application.

Article 2

1. Each State Party shall take effective legislative, administrative, judicial or other measures to prevent acts of torture in any territory under its jurisdiction.

2. No exceptional circumstances whatsoever, whether a state of war or a threat of war, internal political instability or any other public emergency, may be invoked as a justification of torture.

4. An order from a superior officer or a public authority may not be invoked as a justification of torture.

Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW)
3 September 1981, General Assembly Resolution no. 34/180 U.N. Doc. A/34/46

Article 3

States Parties shall take in all fields, in particular in the political, social, economic and cultural fields, all appropriate measures, including legislation, to ensure the full development and advancement of women, for the purpose of guaranteeing them the exercise and enjoyment of human rights and fundamental freedoms on a basis of equality with men.

African Charter on Human and Peoples' Rights
June 27 1981, OAU Doc. CAB/LEG/67/3 21 I.L.M. 58 (1982)

Article 7

Every individual shall have the right to have his cause heard. This comprises:

- o The right to an appeal to competent national organs against acts of violating his fundamental rights as recognized and guaranteed by conventions, laws, regulations and customs in force;
- o The right to be presumed innocent until proved guilty by a competent court or tribunal;
- o The right to defence, including the right to be defended by counsel of his choice;
- o The right to be tried within a reasonable time by an impartial court or tribunal.

- • No one may be condemned for an act or omission which did not constitute a legally punishable offence at the time it was committed. No penalty may be inflicted for an offence for which no provision was made at the time it was committed. Punishment is personal and can be imposed only on the offender.

American Convention on Human Rights

Adopted at the Inter-American Specialized Conference on Human Rights, San José, Costa Rica, 22 November 1969

Article 8

Right to a Fair Trial

1. Every person has the right to a hearing, with due guarantees and within a reasonable time, by a competent, independent, and impartial tribunal, previously established by law, in the substantiation of any accusation of a criminal nature made against him or for the determination of his rights and obligations of a civil, labor, fiscal, or any other nature.

2. Every person accused of a criminal offense has the right to be presumed innocent so long as his guilt has not been proven according to law. During the proceedings, every person is entitled, with full equality, to the following minimum guarantees:

- a. The right of the accused to be assisted without charge by a translator or interpreter, if he does not understand or does not speak the language of the tribunal or court;
- b. Prior notification in detail to the accused of the charges against him;
- c. Adequate time and means for the preparation of his defense;

d. The right of the accused to defend himself personally or to be assisted by legal counsel of his own choosing, and to communicate freely and privately with his counsel;

e. The inalienable right to be assisted by counsel provided by the state, paid or not as the domestic law provides, if the accused does not defend himself personally or engage his own counsel within the time period established by law;

f. The right of the defense to examine witnesses present in the court and to obtain the appearance, as witnesses, of experts or other persons who may throw light on the facts;

g. The right not to be compelled to be a witness against himself or to plead guilty; and

h. The right to appeal the judgment to a higher court.

3. A confession of guilt by the accused shall be valid only if it is made without coercion of any kind.

4. An accused person acquitted by a nonappealable judgment shall not be subjected to a new trial for the same cause.

5. Criminal proceedings shall be public, except insofar as may be necessary to protect the interests of justice.

European Convention on Human Rights

Convention for the Protection of Human Rights and Fundamental Freedoms CETS No.: 005 Open for signatures Rome 4 XI 1950, Entry into force September 3, 1953.

Article 6
Right to a fair trial
In the determination of his civil rights and obligations or of any criminal charge against him, everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law. Judgment shall be pronounced publicly but the press and public may be excluded from all or part of the trial in the interests of morals, public order or national security in a democratic society, where the interests of juveniles or the protection of the private life of the parties so require, or to the extent strictly necessary in the opinion of the court in special circumstances where publicity would prejudice the interests of justice.

Everyone charged with a criminal offence shall be presumed innocent until proved guilty according to law.

Everyone charged with a criminal offence has the following minimum rights:

to be informed promptly, in a language which he understands and in detail, of the nature and cause of the accusation against him;

to have adequate time and facilities for the preparation of his defence;

to defend himself in person or through legal assistance of his own choosing or, if he has not sufficient means to pay for legal assistance, to be given it free when the interests of justice so require;

to examine or have examined witnesses against him and to obtain the attendance and examination of witnesses on his behalf under the same conditions as witnesses against him;

to have the free assistance of an interpreter if he cannot understand or speak the language used in court.

Other instruments

The Bangalore Principles of Judicial Conduct

Adopted by the Judicial Group on Strengthening Judicial Integrity,
endorsed by UN Economic and Social Council Resolution 2006/23
http://www.unodc.org/pdf/corruption/corruption_judicial_res_e.pdf

Draft Principles of Conduct for Court Personnel
Adopted by the Judicial Group on Strengthening Judicial Integrity
http://www.unodc.org/pdf/corruption/publication_jig4.pdf

Council of Europe documents

on standards for legal professionals, including lawyers, judges, and court clerks
This collection includes the 2003 Opinion of the Council of Europe's Consultative Council of European Judges (Conseil Consultatif de Juges Européens, or CCJE) on the principles and rules governing judges' professional conduct, in particular ethics, incompatible behavior and impartiality

Code of Conduct for Law Enforcement Officials

Adopted by United Nations General Assembly resolution 34/169 of 17 December 1979

Basic Principles on the Role of Lawyers

Adopted by the Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders, Havana, Cuba, 27 August to 7 September 1990

Guidelines on the Role of Prosecutors

Adopted by the Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders, Havana, Cuba, 27 August to 7 September 1990

United Nations Standard Minimum Rules for Non-custodial Measures (The Tokyo Rules)
Adopted by General Assembly resolution 45/110of 14 December 1990

United Nations Standard Minimum Rules for the Administration of Juvenile Justice ("The Beijing Rules")

Adopted by General Assembly resolution 40/33of 29 November 1985

Basic Principles on the Independence of the Judiciary

Adopted by the Seventh United Nations Congress on the Prevention of Crime and the Treatment of Offenders held at Milan from 26 August to 6 September 1985 and endorsed by General Assembly resolutions 40/32 of 29 November 1985 and 40/146 of 13 December 1985

REFERENCES

Anderson, James. [N.d.]. "Corruption in Slovakia: Results of Diagnostic Studies." World Bank and United States Agency for International Development. Washington, DC. http://siteresources.worldbank.org/INTWBIGOVANTCOR/Resources/slovrep44.pdf

_____, and Cheryl Gray. 2006. "Anticorruption in Transition 3". World Bank.

Australian Law Reform Commission. 2000. *Managing Justice: Continuity and Change in the Federal Justice System.* Sydney**:** J.S. McMillan Printing Group.

Blank, Joseph, Martin van der Ende, Bart van Hulst, and Rob Jagtenberg. 2004. "Bench Marking in an International Perspective: An International Comparison of the Mechanisms and Performance of the Judiciary System." Netherlands Council for the Judiciary, Rotterdam, May. http://www.ecorys.nl/publicaties/EE10407rap.pdf

Botero, Juan Carlos, Rafael La Porta, Florencio López-de-Silanes, Andrei Shleifer, and Alexander Volokh. 2003. "Judicial Reform." *The World Bank Observer* 18 (1): 61–84.

Brazil, Poder Judiciário, Tribunal de Justiça do Rio de Janeiro. 2004. "Perfil das Maiores Demandas Judiciais do TJERJ." Rio de Janeiro. August.

Carranza, Elías, Mario Houed, Luís Paulino Mora, and Eugenio Raúl Zafaroni. 1988. *El Preso sin condena en América Latina y el Caribe*. San Jose, Costa Rica: ILANUD.

Castelar Pinheiro, Armando, ed. 2000. *Judiciário e Economia no Brasil*. Sao Paulo: Editora Sumaré.

CEJA (Centro de Estudios de Justicia de las Américas). 2003. *Reporte sobre el Estado de la Justicia en las Américas, 2002–2003*. Santiago, Chile.

_____. 2005. *Reporte sobre el Estado de la Justicia en las Américas, 2004–2005*. Santiago, Chile.

CEPEJ (European Commission for the Efficiency of Justice). 2005. "European Judicial Systems 2002." Council of Europe, Strasbourg. http://www.coe.int/T/DG1/LegalCooperation/CEPEJ/evaluation/default_en.asp

_____. 2006. *European Judicial Systems. Edition 2006 (2004 Data)* Strasbourg: Council of Europe. http://www.coe.int/T/DG1/LegalCooperation/CEPEJ/evaluation/default_en.asp

Cooter, Robert D. 1989. "Issues in Customary Land Law." Institute of National Affairs Discussion Paper 39. Port Moresby, Papua New Guinea. August.

Country data. http://www.country-data.com

Dakolias, Maria. 1999. *Court Performance around the World: A Comparative Perspective.* World Bank Technical Paper 430.

Dethier, Jean-Jacques. 1999. "Governance, Decentralization, and Reform: An Introduction." In Jean-Jacques Dethier, ed., *Governance, Decentralization and Reform in China, India and Russia;* 1–46. Dordrecht, The Netherlands**:** Kluwer Academic Publishers.

Faundez, Julio. 2006. "Should Rule of Law Projects Take Community Justice Institutions Seriously? Perspectives from Latin America." *The World Bank Legal Review,* Vol. 2: *Law, Equity and Development,* ed. Ana Palacio. World Bank, 113–140.

Financial Action Task Force/OECD. 2006. "Financial Action Task Force, AML/CFT Evaluations and Assessments. Handbook for Countries and Assessors." June. http://www.fatf-gafi.org/dataoecd/3/26/36254892.pdf

Genn, Hazel. 1999. *Paths to Justice: What People Do and Think about Going to Law.* Oxford and Portland, OR: Hart Publishing.

Grossman, Joel B., Herbert M. Kritzer, Kristin Bumiller, Austin Sarat, Stephen McDougal, and Richard Miller. 1982. "Dimensions of Institutional Participation: Who Uses the Courts and How?" *The Journal of Politics* 44 (1): 86–114.

Hellman, Joel S., Geraint Jones, and Daniel Kaufmann. 2000. "'Seize the State, Seize the Day': State Capture, Corruption, and Influence in Transition." World Bank Policy Research Working Paper 2444. September. http://www.uh.edu/~pgregory/State%20capture.pdf

Hendley, Kathryn. 1999. "The Effectiveness of Legal Institutions in the Transition Economy in Post-Soviet Russia." In Jean-Jacques Dethier, ed., *Governance, Decentralization and Reform in China, India and Russia*. Dordrecht, The Netherlands**:** Kluwer Academic Publishers, 419–46.

_____, Peter Murrell, and Randi Ryterman. 1999. "Law Works in Russia: The Role of Legal Institutions in the Transactions of Russian Enterprises." University of Wisconsin Law School Legal Studies Research Paper Series, Madison.

Human Rights Watch. 2006. "World Report 2006." New York**.** http://www.hrw.org/advocacy/prisons/americas.htm

International Centre for Prison Studies, University of London. 2005. "Prison Brief for Nigeria." http://www.kcl.ac.uk/depsta/rel/icps/wrldbrief/

International Centre for Prison Studies. 2007. "World Prison Brief.", London. http://www.prisonstudies.org/.

Johnston, Michael. 2005. *Syndromes of Corruption*. Cambridge University Press.

Klitgaard, Robert. 1988. *Controlling Corruption*. Berkeley: University of California Press.

Kritzer, Herbert M. 1983. "The Civil Litigation Research Project: Lessons for Studying the Civil Justice System." In *Proceedings of the Second Workshop on Law and Justice Statistics.* Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics, 30–36.

_____. 1999. "Using Public Opinion to Evaluate Institutional Performance: the Experience with American Courts." http://www1.worlbank.org/publicsector/legal/KritzerPremNote.doc

_____, Kenneth Holland, ed., and Bryant Garth, ed. 2002. *Legal Systems of the World: A Political, Social, and Cultural Encyclopedia*. Santa Barbara, CA: ABC-Clio.

National Center for State Courts (NCSC). "Trial Court Performance Standards and Measurement." http://www.ncsconline.org/D_Research/tcps/index.html

Pacific Legal Information Institute. Pacific Region country legal information databases., http://www.paclii.org/pg/cases/PGSC/2005/1.html

Penal Reform International. 2005. "Index of Good Practices in Reducing Pretrial Detention." http://www.penalreform.org

Rose-Ackerman, Susan. 1999. *Corruption and Government*. Cambridge University Press.

Santos, Alvaro. 2006. "The World Bank's Uses of 'Rule of Law': Promises in Economic Development." In David Trubek and Alvaro Santos, eds., *The New Law and Economic Development: A Critical Appraisal,* Cambridge University Press, 253–300,

Stone, Andrew, Brian Levy, and Ricardo Paredes. 1996. "Public Institutions and Private Transactions: A Comparative Analysis of the Legal and Regulatory Environment for Business Transactions in Brazil and Chile." In *Empirical Studies in Institutional Change*, edited by. Lee J. Alston, Thrainn Eggertsson, and Douglass North, 95–128. Cambridge University Press.

Sunstein, Cass, ed. 2000. *Behavioral Law and Economics*. Cambridge University Press.

South African Department of Correctional Services. 2006. Statistics. September. http://www.dcs.gov.za

UK Home Office. 2005. "Population in Custody." Quarterly Brief: April-June. England and Wales. http://www.homeoffice.gov.uk/rds/pdfs05/prisq205.pdf

UNODC (United Nations Organization on Drugs and Crime). 2006. "Assessment of the Integrity and Capacity of the Justice System in Three Nigerian States." Technical Assessment Report. Vienna http://www.unodc.org/pdf/corruption/publications_nigeria_assessment.pdf

Van Velthoven, B.C.J., and M. ter Voert. 2004. *Geschilbeslechtingsdelta 2003* (*Dispute Resolution Delta 2003*), with a summary in English. The Hague: Wetenschappelijk Onderzoek en Documentatiecentrum van het Ministerie van Justitie (Research and Documentation Centre, or WODC, Ministry of Justice).

Wilson, James Q. 2004. *Crime and Public Policy.* In *Crime: Public Policies for Crime Control,* edited by James Q. Wilson and Joan Petersilia, 546–47. San Francisco, CA: Institute for Contemporary Studies.

World Bank. 2002a. "The *Juicio Ejecutivo Mercantil* in the Federal District Courts of Mexico: A Study of the Uses and Users of Justice and Their Implications for Judicial Reform." Report No. 22635-ME.

_____. 2002b. "Legal and Judicial Capacity Building Project (Case Management and Court Administration): Diagnostic Report."

_____. 2003a. "An Analysis of Court Users and Uses in Two Latin American Countries." Report No. 26966.

_____. 2003b. "Brazil: Judicial Performance and Private Sector Impacts." Report No. 26261-BR.

_____. 2003c. "A User's Guide to Poverty and Social Impact Analysis." http://povlibrary.worldbank.org/files/14520_PSIA_Users_Guide/

_____. 2005a. Bank-Wide Legal Modernization Initiative Justice Sector Action Plan. Draft. On file with authors.

_____. 2005b. "Making Justice Count: Measuring and Improving Judicial Performance in Brazil." Report No. 32789-BR.

_____. 2006a. "Doing Business 2007: How to Reform." World Bank/International Finance Corporation.

_____,/International Monetary Fund. 2005. "Financial Sector Assessment: A Handbook." September. http://www.imf.org/external/pubs/ft/fsa/eng/index.htm