Users' Guide For

MEASURING
PUBLIC ADMINISTRATION
PERFORMANCE

UNDP

A Users' Guide to Measuring
Public Administration Performance

A USERS' GUIDE TO MEASURING PUBLIC ADMINISTRATION PERFORMANCE

Cover design and lay-out by Keen Media Co., Ltd.

FIRST EDITION – June 2009

AUTHORS

Kevin Brown
Sarah Repucci

EDITORS

Marie Laberge and Joachim Nahem, UNDP Oslo Governance Centre

ACKNOWLEDGEMENTS

LIST OF INTERVIEWEES

Elissar Sarrouh	IDRC, Canada
Benjamin Allen	Chief of Party, Philippines Governance Project (USAID)
Abdul-Nashiru Issahaku	African Development Bank
John Short	Consultant, Public Expenditure Management
Clay Wescott	Consultant, Public Administration Reform
Christiane Arndt	Researcher, Maastricht School of Government
Guillermo Cejudo	Researcher, CIDE, Mexico City
Shabbir Cheema	East West Center (formerly UNDESA)
Louise Boulter	Senior Lecturer, University of Leicester
Jeremy Clarke	Consultant (formerly DFID Governance Adviser)
Gary Reid	Lead on "Actionable Governance indicators", World Bank
Peter Reed	Consultant, Institutional Development
Tom Christensen	Professor, University of Oslo
Nicola Smithers	World Bank (formerly PEFA Secretariat)
Stephen Sharples	Governance Adviser, DFID
Pamela Gomez	CBPP, Washington DC
Harry Garnett	Consultant (formerly World Bank PSM specialist)
Mike Stevens	Consultant (formerly World Bank PSM specialist)
Gerardo Munck	University of Southern California
Nick Manning	World Bank

CONTENTS

INTRODUCTION

The origins, purpose and structure of this guide

This Guide is intended to respond to an increasing demand from UNDP Country Offices and a wide range of national stakeholders for guidance on the multiplicity of tools and methods that are being used to measure, assess and monitor the performance of public administration (PAP)[1]. There is also a growing demand for more operational and nationally-owned measurement tools for public administration (PA), which is partly a reaction to traditional public administration reforms (PAR) and assessments that were mainly donor-driven.

In this Guide, those tools which produce information and ratings for countries are called "**information sources**" and those which users can apply themselves are referred to as "**assessment tools.**"[2]

The proliferation of different assessment tools has resulted in some confusion about which tools to use, how they should be applied and their weaknesses and strengths. It is our hope that this Guide will help clear up some of this confusion by providing practical guidance to those who are seeking out better ways to measure PA at the country level, including government officials, donor agency staff, reform practitioners, civil society, and researchers. More specifically, readers will find answers to such questions as:

- How to select amongst the existing (or decide to design a new set of) PA indicators?
- How to deal with the preparation and launching of an assessment?
- How to secure national ownership of an assessment process in order to ensure that results are both useful and used by national actors?
- How to ensure that the assessment is rigorous and methodologically sound?
- What to do with the results?
- How to address problems of sustainability

The Guide is essentially made up of two parts. The **first part** critically reviews the existing assessment tools and information sources which are readily accessible to potential users on the internet. It then sets out some practical guidance for users, partly in the form of short stories which illustrate some common measurement problems users may face and how they can be solved. The guidance is based on direct feedback from users of assessment tools, and from a distillation of good practices. To this end, 20 telephone interviews were conducted with PA practitioners (mainly intergovernmental agency staff, consultants and researchers) who have been directly involved in public administration reforms at the country level.

The **second part** is the Source Guide, which is an inventory of extant assessment tools and methodologies. At present, there is no resource that offers a global overview bringing together all extant approaches, tools and methods in the area of PA. The Source Guide is structured in a way to provide detailed information on each tool including: history, objectives, measurement focus, types of information generated, methodology used, strengths and weaknesses (including gender and poverty focus), and website from which a user can access the tool. The purpose of compiling and organising this information is to provide stakeholders engaged in public administration with a resource that can be drawn on for developing new assessment tools or adapting existing assessment approaches to their specific contexts. *It is important to note that the Guide does not provide a new measurement or assessment methodology nor does it provide a specific blueprint to conduct such assessments.*

---

[1] This Guide offers measurement guidance related to the performance of the Public Administration (PAP). It will refer to Public Administration (PA) in general (with the understanding that the focus is on public administration *performance*) and to Public Administration Reform (PAR) when it explicitly deals with issues of reform.
[2] This distinction is based on current practice, although in principle information sources could be used to enable national stakeholders to conduct their own assessments.

The Guide is an important component of UNDP's body of guidance on measuring and assessing democratic governance, developed as part of UNDP's Global Programme on Capacity Development for Democratic Governance Assessments and Measurement.[3] This Programme supports nationally owned processes for assessing and measuring democratic governance and aims to facilitate the development of tools which have broad-based national ownership, are pro-poor and gender-sensitive, and are designed to identify governance weaknesses and capacity gaps.

---

[3] http://www.undp.org/oslocentre/flagship/democratic_governance_assessments.html

CHAPTER I
Assessments of Public Administration Performance

We have used the term "Public Administration" (PA) throughout this guide, although the terms "Public Management," "Public Sector Management", or "Civil Service Management" are often used interchangeably. For the purpose of this guide, Public Administration refers to: (1) The *aggregate machinery* (policies, rules, procedures, systems, organizational structures, personnel, etc.) funded by the state (or local government) budget and in charge of the management and direction of the affairs of the executive (and local) government(s), and its (their) interaction with other stakeholders in the state, society and external environment at national and sub-national levels, and (2) The *management* and implementation of the whole set of (local and national) government activities dealing with the implementation of laws, regulations and decisions of the government(s) and the management related to the provision of public services at national and sub-national levels.[4]

Simply put, public administration reforms can be said to consist of "deliberate changes to the structures and processes of public sector organizations with the objective of getting them... to run better."[5] Depending on the context, it includes mechanisms to improve policymaking and coordination, the building of robust organisational structures, deconcentration and devolution, human resource management, as well as communication and information systems. Particular changes to the physionomy and modus operandi of the public administration will of course be informed by particular ideas and ideologies (such as new Public Management) and they will be shaped by the views and priorities of politicians and government.

Historically, developing countries have tended to follow a similar reform path to the one adopted in advanced countries that have reformed their public administrations. In Africa, for example, reform programmes can be characterised as following three main phases which in practice have overlapped. In the first phase (1980s), reform programmes were launched in the main to achieve fiscal stability by reducing the costs of the bureaucracy. Emphasis was placed on rationalising the payroll, controlling establishments and often outright downsizing of the public service. In the second phase (1990s), once fiscal stability had been more or less achieved, the objectives shifted towards efficiency and effectiveness. This was the era where governments undertook to restructure their public services, to decentralize services and to reform systems of financial and human resource management. The term "building capacity" gained wide currency during this period. In the new century, the emphasis shifted once again with the advent of sectoral programmes (and related sector-wide approaches) which demanded that the public service concentrate on delivering better and more equitable services for its citizens. This new paradigm shift towards 'open government'[6] led to new reforms designed to strengthen the service delivery capacity of the public service and its responsiveness to citizens, with a particular emphasis on target setting.

However, there is no neat dividing line between these phases and governments have carried forward unfinished agendas as new programmes have been initiated. As a result, many developing country governments find themselves focusing on many objectives simultaneously and as a result carrying out very broad programmes of PAR. Until recently, having a "comprehensive" programme was considered "good practice" because successful reform usually requires tackling the whole system rather than isolated parts of it. But today big bang approaches are increasingly being questioned, as positive results with comprehensive reforms remain scarce. While certain reforms undoubtedly need to be steered at the corporate level, sectoral approaches are getting increasing attention both from the governments and from the donor community. Hence, not all governments are willing or capable of pursuing comprehensive reform agendas.

---

4   UNDP Practice Note on Public Administration Reform, 2003
5   Christopher Pollitt and Geert Bouckaert, Public Management Reform: A Comparative Analysis, Oxford University Press, 2000.
6   This paradigm shift towards 'open government' was inspired by the World Public Sector Reform Report, "Unleashing capacities for public sector management".

1.1   Our classification of assessment tools and information sources

This guide starts by examining the key elements that need to b taken into account when measuring the performance of the public administration[7], in terms of their contribution to the efficiency, quality and effectiveness of public administration. The model can be applied when examining the performance of the public administration at the aggregate level, the sectoral level, or the organisational/agency level. Our working typology is summarised in Box 1 below.[8] Some of these themes will be familiar to readers, others less so. It is therefore necessary to make the following clarifications:

*   Civil service leadership has been added as a theme in recognition of the critical role that senior managers play, in some cases as part of a separate executive service, both in driving reforms and maintaining an effective public service;

*   Service delivery is regarded by many commentators as an outcome of reform, where public administration engages with the citizen. But it is also a key driver in that ambitious poverty reduction programmes require a more efficient and effective public administration in order to succeed. As such, improving the processes and capacities for delivering public services is an increasingly significant focus in many reform programmes.

BOX 1 – A WORKING TYPOLOGY FOR MEASURING PUBLIC ADMINISTRATION PERFORMANCE

*   **Civil service management:** legal and ethical framework, HRM policies and procedures, institutional framework
*   **Public financial management:** budget preparation and execution, accounting systems, audit and legislative scrutiny
*   **Government policy making:** processes, structures and capacity for analysing problems, identifying and appraising options, consulting with stakeholders, decision-making, monitoring and evaluation
*   **Leadership:** selecting, rewarding, deploying and developing capacity of senior public servants
*   **Service delivery:** identifying client needs, developing service standards and targets, monitoring performance, building capacity to deliver

---

1.2   Balancing normative and technocratic assessments

This classification is technocratic because it focuses on the content of particular reform ingredients rather than the quality of public administration per se. Further, it ignores both the political environment in which PAR is implemented and the specific aims which are sought by the public administration. The paradigm shift towards more "open government" as an essential component of a democratic governance infrastructure has widened the debate on PAR to also include the fostering of dynamic partnerships with civil society and the private sector, and ensuring broader participation of citizens in decision-making and public service performance monitoring. It means that apart from a critical focus on delivering services, special attention is also needed to transform the public service into a representative and accountable institution.

UNDP sees the public administration not only as an instrument to ensure effective policy making, sound management of resources and the delivery of essential services, but also as an essential means to advance the international norms and principles that underpin the democratic governance paradigm in a given society. For instance, rule of law deficits in the public administration – lack of written decisions, or resistance of lower agencies to comply with decisions and rulings – can have a serious impact on individuals and communities. By situating the public administration within a rule of law framework, the 'users' of the system become rights-holders, able to legally claim services and hold state agents accountable.

---

[7]   For the purposes of this guide, the public administration will refer to the executive branch of the central government. It includes the ministries which serve the executive, as well as non-ministerial departments, semi-autonomous bodies and front-line service delivery units such as schools and hospitals.

[8]   A further component, the machinery of government (which is concerned with the rules, institutions, and structure of the administration necessary to carry out government policy), has been omitted from this classification because we were unable to find a single tool which purported to measure it. This is a major gap in the range of assessment tools because restructuring is a typically a key component of PAR in developing countries. Finally, procurement systems and E-government are not included as separate elements of PAR in our typology.

Another example of how a democratic governance framework calls attention to particular dimensions of the public administration relates to the principle of accountability. Still today, in many bureaucracies around the world, accountability is essentially internal, through a hierarchy of supervisors and horizontal control structures, with little concern for accountability to the public or to elected representatives. Client charters are easily applicable methods of downward accountability. Public hearings and public grievance mechanisms that allow citizens to report abuses of authority can also promote democratic oversight at the local level.[9]

Clearly, the normative framework selected will influence the choice of interventions, the priorities within them and how they should be implemented in particular country contexts. It will also affect the particular themes and indicators which are considered relevant in any assessment exercise. For instance, the Country Governance Assessment framework, developed by the Asian Development Bank, provides guidance questions for the evaluation of revenue administration and public expenditure management. While many of these questions are purely technical, some are informed by democratic governance principles, such as responsiveness ("*Do budget priorities reflect underlying needs and demands of civil society?*") and transparency ("*Is the budget published and made available to the public?*") In later sections of this guide, we look at more concrete examples where the choice of indicators reflects a democratic governance perspective of public administration.

1.3   Scope of information sources and assessment tools

With the exception of public financial management, public administration reform has lost some momentum in recent years as a consequence of the explosion of interest, for good reasons, in the broader concept of governance among the donor community. In response to this, many information sources and assessment tools have emerged in recent years for measuring governance and corruption[10], but not that many measure different aspects of public administration performance, even though several anti-corruption assessment tools relate directly to PA. However, there is a new realisation in some quarters, notably the World Bank and the OECD, that there is a gap in the measurement of PA and its impact on the effectiveness of government that needs to be filled. As a result promising new tools are now being developed.[11] The recent financial crisis and its impact on the performance of governments and the delivery of services in particular have put public administration high on the governance agenda again

There are far more assessment tools available to users than information sources

Our research has unearthed a number of relevant existing tools and information sources which can be used to measure different components of PAP. They are all publicly accessible, although a few require a simple registration process before they can be used. No licensed instruments, which can only be administered by a particular organization on the payment of a fee, have been included. Of the 25 instruments highlighted in Table 1, 18 are assessment tools and only 7 are information sources.[12]

There is a reasonably good range of assessment tools which are accessible to users if we search beyond the international aid community

These 25 instruments can be further categorised as follows:

- Purpose-built tools which purport to measure a particular aspect of public administration performance
- Broader governance assessment tools and information sources which include a component on public administration
- Tools designed for a different purpose, but which measure the relevant processes

---

9   A UN Discussion paper on the approach to public administration, local governance and financial transparency and accountability in post-conflict peace building operations (draft), Dec. 2008
10   See UNDP, Governance Indicators: A Users' Guide, second edition (2007)
11   Such forthcoming tools include the World Bank's HRM Actionable Indicators and the OECD's 'Government at a Glance' indicators.
12   A number of broader governance tools are excluded from our list because their measures of public administration are limited.

The inclusion of the third category provides the potential user with a significant amount of choice. For instance, the Common Assessment Framework (CAF) developed by the European Institute of Public Administration was originally designed to assess the effectiveness of public sector organizations (from a Total Quality Management perspective) but it can be adapted to measure systemic issues. And there are an additional two tools (the World Bank's "Actionable Governance Indicators" and the OECD's "Government at a Glance."

As a first screen for users, Table 1 indicates the thematic areas covered by each assessment tool and information source.

Table 1 – What do assessment tools and information sources measure?

Ref	Source	Name of assessment tool / Information source	Leadership	Civil service	Public financial management	Policy making	Service delivery
Assessment tools						
1	World Bank	Quantitative Service Delivery Surveys					X
2	World Bank	Public Expenditur Tracking Surveys			X		
3	Public Affairs Centre	Citizen report cards					X
4	Asian Development Bank (ADB)	Country Governance Assessment		X	X	X	
5	European Institute of Public Administration (EIPA)	Common Assessment Framework	X	X	X	X	X
6	UK Civil Service	Capability reviews	X				
7	UK Civil Service	Self-Assessment Tool for Customer Service Excellence					X
8	Multi-institution project (World Bank, IMF, EC, DFID, SPA)	Public Expenditure and Financial Accountability (PEFA)			X		
9	World Bank – Netherlands Partnership Program (BNPP)	Public Officials' Survey		X	X		
10	UNDP	Country Assessment in Accountability and Transparency (CONTACT)			X		
11	World Bank	Evaluation Matrix of Civil Service Human Resource Management in the EU		X		X	
12	SIGMA (OECD & EU)	Control & Management System Baselines		X	X	X	
13	State of Texas	Human Resources Self-Assessment Guide		X			
14	Management Sciences for Health	Human Resource Management (HRM) Assessment Instrument		X			
15	Inter-American Development Bank (IADB)	Analytical Framework for Institutional Assessment of Civil Service Systems		X			
16	Commonwealth Secretariat	Engendering Budgets: A Practitioners' Guide to Understanding and Implementing Gender-Responsive Budgets			X		
17	Transparency International	National Integrity Systems	X	X			
18	World Bank	HRM Actionable Governance Indicators (*currently under development*)		X			
19	World Bank	Diagnostic Framework for Revenue Administration			X		

Table 1 – What do assessment tools and information sources measure? (continued)

Ref	Source	Name of assessment tool / Information source	Leadership	Civil service	Public financial management	Policy making	Service delivery
Information sources							
20	OECD	Government at a Glance (*currently under development*)		X	X	X	X
21	Bertelsmann Siftung	Management Index, Bertelsmann Reform Index	X	X			
22	International Budget Project	Open Budget Initiative			X		
23	World Bank	Cross-National Data on Government Employment & Wages		X			
24	IMD	World Competitiveness Yearbook, Government Efficiency			X	X	
25	World Bank	Governance Matters, Government Effectiveness		X		X	X
26	World Bank	Country Policy and Institutional Assessment (CPIA) Criteria 13: Quality of Budgetary & Financial Management Criteria 15: Quality of Public Administration		X	X	X	X
27	Global Integrity	Integrity Indicators	X	X	X	X	

There is one assessment tool, EIPA's Common Assessment Framework (CAF), which attempts to measure all five dimensions of PAP. This assessment focuses on the processes and results of an organization, rather than the public administration environment in which it sits. However, because organizational processes will reflect the constraints imposed by the central rules and regulations, application of the tool across a number of public institutions can generate evidence of systemic PA problems. In addition, there are two tools which measure three of the core themes, namely human resource management, public financial management and policy making. These are the Asian Development Bank's Country Governance Assessment and SIGMA's Control and Management System Baselines.

There is a preponderance of assessment tools and information sources which measure public financial management reform and human resource management reform

When looking at both assessment tools and information sources, we find 13 instruments measuring PFM and 15 instruments measuring HRM. Both are critical reform areas. However, PFM tools have attracted a greater interest than HRM tools within the development community in the last few years. As a consequence the PFM tools tend to be more contemporary than the HRM tools. In part, this situation reflects the self-interest of donors in the effectiveness of their aid, as well the shift in aid practice towards direct budget support mechanisms which give rise to concerns about fiduciary risk. It also reflects the professional biases of international agencies, which tend to be dominated by economists and financial management specialists. But traditionally also, these two sectors were seen as the main drivers of public sector efficiency (financial and human resources). Attention to leadership, service delivery, the quality of policy making and information and communication capacity have come at a later stage. However, the forthcoming Actionable Governance Indicators from the World Bank will to some extent restore a balance.

Instruments which focus on a single theme can generate deeper assessments than the broader tools

Tools which attempt to measure a number of themes, in particular the broader governance assessments, do not generally measure them in great depth. In contrast, instruments which focus on a single theme provide for a much deeper assessment of the phenomena being measured in a relatively short period of time. Our mapping has identified the following tools and sources which focus on a single theme (Table 2):

Table 2 – Assessment tools and information sources which focus on a single theme

| Civil Service Management	• Human Resources Self-Assessment Guide • Human Resource Management Assessment Instrument • Cross-National Data on Government Employment & Wages • Analytical Framework for Institutional Assessment of Civil Service Systems
Public Financial Management	• Public Expenditure Tracking Surveys • CONTACT • Public Expenditure and Financial Accountability • Engendering Budgets: A Practitioners' Guide to Understanding and Implementing Gender-Responsive Budgets • Open Budget Initiative • Diagnostic Framework for Revenue Administration
Service delivery	• Quantitative Service Delivery Surveys • Citizen Report Cards • Self-Assessment Tool for Customer Service Excellence
Leadership	• Capability Reviews

1.4   Designing a PAP assessment

The majority of instruments have been designed by international agencies, which means that special efforts must be made to achieve host government ownership of the findings

More than 60 per cent of the information sources and assessment tools captured have been designed (or are in the process of being designed) by official international agencies. Others have been developed by public sector organizations in advanced countries and a private consulting company. Few instruments have been developed by civil society organizations – NGOs, academic and research organizations.

In practice, therefore, a national government wishing to carry out a PAP assessment will often be choosing a tool which has been designed by an international agency for application in a number of countries. Users will therefore have legitimate concerns about the relevance of the indicators and methodology to their particular situation and as a result the findings may lack face validity. Potential users are therefore encouraged to examine the procedure that was used to develop the instrument. For example, to what extent was a sample of national stakeholder consulted in the design? Or, was the tool first tested in developing country contexts and subsequently refined before its wider application?

PAP assessments are only useful to the extent that they lead to action by national stakeholders. And action will only occur when findings are perceived as valid by the user, which in turn will depend on "ownership" of the assessment process. Ownership is therefore a central theme in this guide. Our contention is that ownership is more likely to be achieved if the user is able to make a contribution to the design of the assessment tool. As illustrated in Figure 1, the design involves three key steps: deciding upon objectives, selecting specific measures and developing the methodology. All three steps are important for generating accurate and valid data for users, and for building ownership of the assessment process.

Figure 1 – Users' contribution to the design of the assessment is key for securing ownership
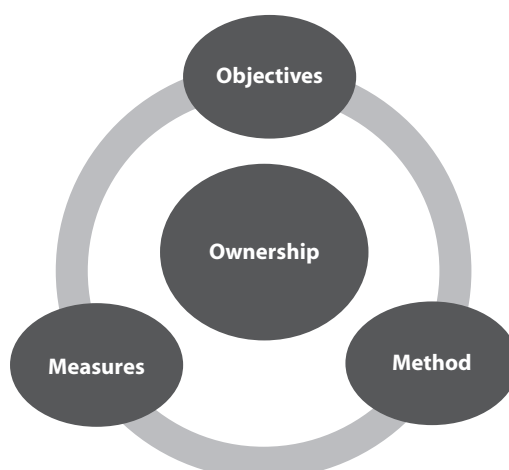
1.5   What are the objectives of the user?

An essential criterion for selecting a source or tool is whether it meets the objectives of the user. By objectives, we mean the ultimate purposes for which the information will be utilised. As a very first step, users should have a clear understanding of why the assessment is important, why they are carrying it out. In browsing through the official information provided by the owners of different tools, the potential user will find a veritable dictionary of words used to describe the intended objectives. For example, the list includes assessment, benchmarking, cross-country comparison, research, dialogue, diagnosis, improving performance, aid allocation, feedback, tracking changes, measuring impact of reform and so on. Some tools are not explicit about objectives and simply refer to the information and results that the source or tool is capable of generating. When this is the case, clarity of purpose for undertaking an assessment is all the more important if users are to be able to select the right tool only by examining the information generated by it.

The user should be absolutely clear on his or her objectives for the assessment exercise before choosing a tool

For this guide, objectives have been classified into five categories based on the information provided in the official literature on the tool and, in some cases, our interviews:

- *Diagnosis & planning:* where the information generated constitutes a formal review of strengths and weaknesses feeding into a planning exercise in which decisions are made on reform priorities. Many tools claim to be useful for diagnostic purposes. The reality, however, is that the majority provide data on high level indicators corresponding to the immediate objectives of reform which need to be analysed further. Such tools can identify the weak areas for further investigation as part of a more intensive diagnostic process. [Example – HRM self-assessment]

- *Monitoring and accountability:* in which information paints a picture of how well government is performing in a particular thematic area. Monitoring is useful to establish the current situation in relation to an established standard against which future improvements can be measured. Many tools can be used for monitoring purposes. But if the user wishes to monitor progress of an ongoing reform programme, the measures selected will need to be customised to, reflect its strategic objectives [Example – Citizen Report Cards]

- *Cross-comparison and benchmarking:* in which information is used to compare one country with another, usually by calculating a score for the phenomena being measured. Benchmarking is a further development which involves a measurement against a standard, or "good practice." This is a common objective, particularly for those sources and tools which originate in the donor community. [Example – Cross-National Data on Government Employment and Wages*]*

- *Dialogue and joint decision-making:* where information can be used to initiate a dialogue about changes and improvements. Dialogue is a modest objective which is particularly relevant where donor-government partnerships are being promoted in line with the Paris Declaration. Dialogue can form the first step in a more in-depth diagnosis [Example – Open Budget Index]

- *Resource allocation:* in which the outcome of the exercise is linked to the allocation of aid. This is not an explicit objective of any of our tools or sources, but it is included because it may be a perceived objective by national governments [Example – PEFA]

The stated objectives of sources and tools are expressed in terms of this classification in Table 3. International agencies are interested in measurement for a variety of purposes – dialogue, diagnostic, monitoring and international benchmarking. In contrast, national governments are mainly interested in measuring PA for diagnostic and monitoring purposes, as well as for priority setting, and to a lesser extent for dialogue and joint decision making. Middle income countries aspiring to become developed nations are often interested in international benchmarking, but most developing countries less so.

Many of the sources and tools are serving multiple objectives. In general, where a single tool has many objectives it tends to be difficult to meet them all satisfactorily. Some objectives are complementary. For instance, international comparisons and benchmarking can form part of a locally applied diagnostic tool or monitoring exercise. And information for dialogue can be useful for diagnostic purposes. However, information for dialogue is not the same as a diagnostic, which examines causes as well as symptoms. If the user has multiple objectives, it will be necessary to "mix and match" different tools.

Assessment exercises usually work best when the user has a single objective, and where this mirrors the objectives of other stakeholders

It is often claimed that a tool can be used for both diagnostic, and for monitoring and accountability purposes. This is not strictly correct. A diagnostic tool designed to formulate a reform plan will normally need to delve deeply into relevant processes (e.g. recruitment and selection of civil servants) to determine how they are working and why they are working the way they are. This level of detail may be required for monitoring purposes by the specialists tasked with making the changes. But it will be less useful for reform managers who are more likely to want information on the immediate results which the process changes are intended to produce. In practice, therefore, the information needs for diagnosis and monitoring will differ, although there will be areas of overlap.

Ideally, there should be agreement on the objectives of the exercise among all stakeholders. Objectives should be stated explicitly. Problems can arise when the user perceives that a key stakeholder has a hidden agenda. An example is PEFA is intended to monitor the performance of a country's public financial management system, as well as providing a platform for dialogue between donors and governments. However, because of the comprehensiveness of the tool, some assessments have formed part of a fiduciary risk assessment. In such cases, a national government may attempt to cover up any deficiencies because this could reduce the prospects of incremental development aid. This potential blurring of objectives occurs in large measure because of the success of PEFA which has tended to supersede many former tools. The problem can be overcome through extensive consultation between government stakeholders and the assessment team during the process, and by placing emphasis on strengthening PFM systems over time, rather than meeting a fixed standard.

Objectives, both explicit and perceived, can influence behaviour and hence the quality and reliability of the information generated

Where the objective of the user is to provide information for monitoring and accountability purposes, there will be a tendency for those providing the information either to set the target low or to inflate the measure so that performance appears better than it is. A diagnostic tool can lead to covered up problems where those providing the information have a vested interest in preserving the status quo. This will occur when there is a separation between the user, such as a Minister or senior government official, and the respondents, who may be employees at the operational level.

1.6   What do the tools and information sources actually measure?

A second criterion for choosing a tool is the relevance and quality of the information generated. After the user has identified the range of instruments which meets his objectives, he should examine more closely at the indicator set they employ.

There is a critical distinction between practices "de jure" (in law) and those "de facto" (in use.) Often, official practices are not carried through or implemented only partially. For this reason most tools attempt to measure practices "in use" although this is not always explicit in the way the indicator or question is framed.[13] The World Bank's forthcoming HRM Actionable Governance Indicators will classify each question as "in law" or "in practice." There is always a risk that respondents will answer the easier question (whether a written procedure exists), especially when there is no clear cut answer on whether it is implemented in practice. As a rule, when an assessment enquires whether a written procedure exists, it should also ask the extent to which it is actually applied.

The user has choices in the thematic areas of service delivery and human resource management reform, where there is significant diversity in indicator sets and methodologies. The tools measuring PFM and government policy making are more uniform, in part reflecting a consensus among agencies on "good practice."

---

[13]   The exceptions are the SIGMA Control and Management System baselines and the Matrix for EU Accession which place considerable emphasis on the existence of policies and procedures in law.   In these tools the distinction between practices "de jure" and those in use is made explicit.

Service delivery

There are four service delivery tools presented in our source guide, which measure three different aspects of service delivery. The World Bank surveys tend to focus on concrete dimensions of service delivery – such as quality, timeliness, access and efficiency – which can be tailored to the sector or organization. Usually, quantitative information is obtained from the service providers' records. Both the Citizen Report Card and the EIPA's CAF measure service delivery, but from the perspective of the user rather than the provider. The former tool actually asks citizens for their views whereas the latter simply requires management to provide an opinion on customer satisfaction.

In contrast to these two tools, the UK Civil Service Self-Assessment Tool for Customer Service Excellence emphasizes how services are provided, rather than what is actually delivered to the customer. The focus is on improving processes and capacity that are likely to deliver value to the customer. It therefore looks at the determinants of customer satisfaction. None of these tools are necessarily superior to any other. Each can be used as a stand-alone tool or they can be used in combination to provide a more comprehensive picture of service delivery.

Public Human Resource Management

In measuring HRM reform the user has a more difficult choice to make. There are nine assessment tools in our source guide, one which is under development (the World Bank's HRM Actionable Governance Indicators). Three assessment tools – authored by EIPA, MSH, State of Texas – measure policies and processes covering the full range of human resource management activity. In addition, the MSH tool examines the capacity of the organization to manage its people. These tools all derive their indicators from good private sector practice which are arguably applicable to public sector organizations.   Indeed, they provide a more comprehensive coverage of HR practices than the purpose-built tools.

However, there are some important differences between HRM in the public and private sectors. The first is that in the public service HR practices are usually governed by a Civil Service Law which defines the basic principles of employment for all public service bodies, with procedural details established in subsidiary regulations[14]. A second difference is that employees in the public service are required to uphold certain ethical and behavioural standards, such as impartiality, which are derived from a common set of values. A final critical difference is that, because the principle of entry to the service on merit must be upheld, special arrangements are necessary to check that the executive does not abuse its power to make appointments for reasons of nepotism or political patronage. Independent oversight of the executive can be achieved either by establishing a Public Service Commission or an Ombudsman Office reporting to Parliament or providing legal mechanisms for individuals to seek redress through the Courts.

There are three assessment tools which recognise the distinctive nature of human resource management for the civil service. These are the World Bank's Evaluation Matrix of Civil Service HRM in the EU, the SIGMA's Control and Management System Baselines, and the Asian Development Bank's Country Governance Assessment tool. All of these pay attention to the legal framework, the institutional arrangements for managing the civil service, and the conduct of civil servants. At the same time, they also cover the key HRM policies and procedures. These are all good tools in terms of the breadth of indicators, but the World Bank tool is arguably the most comprehensive. The World Bank's forthcoming Actionable Governance Indicators for civil service management promises to be an improved version of the EU accession instrument.

Public financial management

Public financial management is well trodden territory and the user should not expect to find much variation in the indicators employed in the many tools available. The technical scope of PFM embraces budgeting, accounting, audit and legislative oversight. The PEFA tool is the most contemporary tool which is well researched and supported by a consortium of international agencies. The framework of six dimensions and 31 high level indicators is comprehensive, reflecting the problems commonly found in developing country systems, and also many developed countries[15]. The Common Assessment Framework (CAF) by EIPA is also comprehensive in scope, but it tends to view good practice from an organizational rather than a systems perspective. Both of these tools can be recommended to potential users. UNDP's CONTACT is a more detailed question-based instrument which finance specialists will find useful.

---

[14] While private companies also have similar statutory frameworks and implementation arrangements, they tend to differ from one company to another, as opposed to Civil Service Law which establishes a common framework for all public service bodies.

[15] The Government of Norway, for example, did a self-assessment using the PEFA instrument and did not perform well on a number of indicators. This indicates the somewhat ambitious nature of PEFA standards, perhaps too ambitious for many developing countries.

The Open Budget Index (OBI) is an information source which complements the more technical tools by assessing budget transparency. It provides information on the availability of budget documentation, the executive's budget proposal and the budget process in 59 countries. OBI has built into its indicator set many of the key elements of the budget process that are measured by PEFA. For example, in order for a country to score highly on the OBI index, its budget must be comprehensive, it must adopt a standard budget classification and resources reflected in the budget must be aligned with government's policy objectives. These processes guarantee that the information provided to civil society is meaningful.

Government policy making

The policy making assessment tools are all checklists which capture very similar information. They focus on the policy making process, but they also examine capacity to some extent. Of these, the SIGMA checklist is perhaps the most widely accepted and recognised. It examines: (a) the alignment of sectoral policies with overall government strategy; (b) rules for the preparation and submission of policies; (c) provision for inter-ministerial consultations; (d) the efficiency and quality of decision making; and (e) capacity for preparing and negotiating policies in ministries. The CAF tool differs from the others in that it measures the development of strategy and policy from the perspective of individual public organizatiions rather than the wider government policy process.

Leadership

The two leadership assessment tools both measure the behaviour of senior managers, although there are differences in the competences selected. One important difference is that the CAF assesses relationships with stakeholders, including with politicians, a dimension which is missed by the Capability Reviews.

1.7   What is the normative basis of these assessments?

Governance assessments are concerned with the extent to which particular principles and values (e.g. transparency, accountability, participation, integrity) are put into practice. They are therefore inherently normative in their orientation. In contrast, PAP assessment tools tend to focus on practices rather than principles which often obscure any value judgements from view. In addition, they are typically neutral with respect to the policy objectives (e.g. poverty reduction, gender equality) that are sought from better public administration

The available tools for PFM and Government policy making incorporate indicators that are universally accepted

Is there is a set of good practice indicators for each thematic area that are universally valid, or do the indicators inevitably reflect the values and biases of the tool's author(s)? PEFA was a genuine attempt by a number of donor agencies to develop a universal set of high level PFM indicators relevant to all countries. On the whole the assessment tool meets this ambitious claim, given its successful application in many different countries.[16] A similar claim can also be made for the OECD/SIGMA policy making tool which has been embraced by many practitioners. Its standards are internationally valid, not just for European countries and those aspiring to join the EU, for whom they were originally intended.

One of the reasons that these tools have acquired this status is they have been kept intentionally simple and restricted to a limited number of high level indicators, which correspond to broad and widely accepted standards. The beauty of the OECD/SIGMA tool is that it does not prescribe the precise ways in which these standards should be accomplished. For instance, one standard states that: "effective inter-ministerial cooperation should take place… through the exchange of information," but it does not say what mechanisms should be used to secure cooperation or what media should be used to exchange information. These decisions are left to the discretion of the government. As soon as one starts to prescribe the specific procedures for meeting agreed standards, then legitimate choices start to emerge and the tool will relinquish its claim to universality.

---

[16]   One practitioner indicated that the indicator on parliamentary scrutiny was geared towards the Anglophone public administration tradition. However, this is only one indicator out of 31.

But, as yet, there is no equivalent public human resource management assessment tool which can claim to embody a universal set of indicators

There are a number of tools, which have a number of indicators in common, but also many divergent measures. It appears that the main reason for the existence of many diverse tools is that no agency has yet invested in the research and development of a universal tool which draws together the good features of the existing tools. The different indicator sets simply reflect the different emphases introduced by the practitioners who developed them.

Both PFM and government policy making have evolved to a point where there exist universal indicator sets that can be legitimately applied to any assessment exercise. The user therefore does not need to "reinvent the wheel." However, in the relatively neglected area of public human resource management service reform, there is far less consensus on what constitute valid indicators. For this thematic area users are advised to customise their assessment instrument, drawing upon the available indicators in different tools.

1.8   Are good governance principles embedded in assessment tools?

Even if such principles are not explicit, it may be that they are reflected in specific indicators. Accordingly, as examples, we examined the extent to which the principles of "responsiveness"[17] and "transparency" are incorporated within the various assessment tools.[18] In the main, such principles are not included, with the exceptions shown in Box 2 below.

BOX 2 – INCLUSION OF RESPONSIVENESS AND TRANSPARENCY PRINCIPLES: EXAMPLES

Country Governance Assessment (Asian Development Bank)
- Participation of the public in the policy-making and legislative process
- Are criteria for selection of public servants transparent?
- Do budget priorities reflect underlying needs and demands of civil society?
- Is the budget published and made available to the public?

Self-Assessment Tool for Customer Service Excellence (UK Civil Service)
- Information and access: accurate and comprehensive information is delivered through the most appropriate channel for customers
- Customer insight: effectively identifying customers and consulting them in a meaningful way

Public Expenditure and Financial Accountability (WB, IMF, EC, DFID, SPA)
- Transparency of inter-governmental fiscal relations
- Public access to key fiscal information

Evaluation Matrix of Civil Service HRM in the EU (World Bank)
- Establishment of citizens' grievance process in law
- Existence of an oversight body to ensure the application of civil service policies
- Public advertisement of posts to ensure equal competition
- Access of civil servants to their performance appraisal
- Disciplinary principles based on transparent and fair principles

Open Budget Index (International Budget Project)
- The entire tool measures the access of citizens to budget information

---

[17]   Defined as taking account of citizens' aspirations and needs
[18]   Other relevant public administration principles include accountability, participation, fairness, integrity

Since the principle of "responsiveness" is inherent in the concept of customer service, it is not surprising that it is captured in the indicators for the customer service excellence tool. It is surely relevant to government policy making too. Yet the existing tools focus on internal processes, in so doing downplaying the interface between government actors and civil society. There is a strong case for including additional measures to reflect this principle. The principle of "transparency" is relevant to the government budget and to certain human resource management processes (e.g. recruitment and selection, performance appraisal, discipline.) And it does seem to be captured reasonably well by the better PFM and public HRM tools.

1.9   How can PAP assessments focus on particular objectives of government policy?

UNDP's normative standpoint in all governance support is a concern for the rights, needs and interests of the poor, women and other vulnerable groups. Can PAP assessments capture such impacts? "Service delivery" is the one thematic area where the tools in our source guide pay explicit attention to the impacts on disadvantaged groups and women.[19] For instance:-

•   Citizen report cards gather the views and perceptions of the actual users of services, who can be stratified into different groups, including poor people and women;

•   The World Bank's Quantitative Service Delivery Survey examines the quality and efficiency of the service provided which, depending on the survey design and location, can also collect information on the beneficiaries of the services; and

•   The UK Civil Service's Self-Assessment Tool for Customer Service Excellence includes a question about the processes in place to identify the needs of "hard to reach" groups.

However, the assessment tools covering the three core themes – public HRM reform, public financial management reform, government policy making – are primarily concerned with the measurement of internal processes which are one-step removed from the service delivery processes which deliver benefits directly to citizens

In these cases it is less obvious how a pro-poor perspective can be captured within the assessment tool itself. Disaggregating indicators by particular target groups (e.g. poor people, women) only works where the direct impact of the concept being measured (e.g. corruption) is experienced by such groups.

The solution lies in linking specific reform interventions to desired policy objectives in particular country contexts

Country Poverty Reduction Support Programmes – which incorporate the MDGs – provide an excellent foundation for this thinking because they specify the high level policy goals which governments wish to pursue. The steps to be taken are as follows:

•   **Step 1.** Select a specific policy goal
•   **Step 2.** Identify the outputs that will contribute to this goal
•   **Step 3.** Develop concrete indicators for these outputs
•   **Step 3.** Carry out the PAP assessment

The output indicators provide the focus for a PAP assessment. A reform programme can then be designed based on the gaps identified between the actual and the desired situation. Box 3 shows examples of concrete reform outputs which are relevant to pro-poor policy objectives.

---

[19]   Significantly, however, none of these tools are designed to measure the extent to which disadvantaged groups have access to services because they focus on existing service users.

1.10 Contextualizing PAP assessment tools

The choice of appropriate indicators will depend on the level of reform maturity

Consider three country types – post-conflict, developing and middle income. Although there will always be exceptions, their levels of maturity in a particular reform area will typically mirror their levels of development. Assessment tools are generally designed to meet the specific requirements of a typical developing country which may not be appropriate for a post-conflict or a middle income country. Countries emerging from conflict will have experienced a collapse of many of the basic systems and rules of public administration. Where fiscal discipline is the most immediate problem that needs to be faced, there seems little purpose in asking questions about medium-term budgeting or the efficiency of spending. Similarly, a middle income country which has aspirations to become a developed country is more likely to want to know whether its HR processes reflect modern practice, than whether its civil service law is sound. For this reason, users should select those indicators that are most appropriate to their level of maturity in a particular thematic area.[21]

Indicators need to be consistent with country traditions of public administration

Countries not only differ in their relative levels of development; they have different traditions of public administration which effectively constrain the choices they can make. The theory of "path dependency" states that history and traditions will affect what reforms are feasible in particular countries. It is no coincidence, for instance, that the developed countries that have been most receptive to "New Public Management" approaches to reform are those with an Anglo-Saxon administrative culture, such as New Zealand, the UK or the USA. In contrast, NPM has had far less impact in countries in Continental Europe which uphold the "Rechtstaat" model of public administration – such as France and Germany – although both have experimented with some elements. Users will need to be open to alternative perspectives, whilst at the same time being willing to challenge those indicators which do not feel right.

Indicator sets should be regarded as flexible frameworks which will need to be reviewed and revised as reforms evolve

It is often claimed that, if an assessment is repeated in the same country at a later date, it will measure reform progress over time. This is a goal to which PEFA aspires. Because changes in the high level indicator set will occur relatively slowly, it is recommended that a PEFA assessment is conducted every two to three years. The assumption behind this claim is that government's reform objectives and trajectory remain unchanged. It is doubtful, however, that this is a wholly valid assumption. Whilst the direction of reform is likely to remain constant, the pathway chosen may change over time. Case studies developed at the World Bank have demonstrated that it is rare in the second year of a reform programme

---

[20]   The solution will depend upon whether a devolved recruitment and selection system is adopted.
[21]   One PEFA practitioner felt that a "gold standard" version could be adopted for countries whose PFM systems are more advanced.

that a government will be undertaking an identical programme to the one planned in the first year. Operational indicators may therefore change where learning demonstrates that a particular approach will not work. Priorities may also evolve in response to changed circumstances and high level indicators may need to be adjusted accordingly. An assessment instrument must accordingly be regarded as a flexible tool, to be adjusted as the reform trajectory evolves.

1.11 Should indicators measure inputs, processes, outputs or outcomes?

This guide follows the approach adopted by the OECD in "Government at a Glance" by defining:-

- Inputs as units of labour, capital, goods and services sacrificed for the production of services (e.g. for health services, inputs can be defined as the time of the medical staff, the drugs, the electricity, etc.)

- Public sector processes as structures, procedures and management arrangements with a broad application within the public sector (e.g. social and financial audit of local clinics)

- Outputs as the services that are delivered to the end user by a public sector organization (e.g. number of babies delivered safely, number of pupils completing primary education);

- Outcomes as the impact of these services on the end user (e.g. reduced maternal mortality, improved income)

Public administration reforms contribute to the achievement of outputs and outcomes, which will typically be reflected in the higher level objectives of the overall reform programme. Reforms create changes in processes (structures, skills, staffing and incentives), which in turn lead to change in outputs.

The OECD's "Government at a Glance" notes that there is a dearth of output indicators in existing sources and tools. In this Guide, only three tools were found to measure outputs exclusively, namely the Quantitative Service Delivery Surveys, the Public Expenditure Tracking Survey and the Citizen Report Card. In addition, the EIPA's CAF measures both outputs and outcomes as one element in the assessment.

Most tools and sources assess processes. For instance, the civil service management component of the EU accession matrix tool measures: (a) the Legal and Ethical Framework; (b) the Institutional Framework (structures); (c) Pay Policy and Management (incentives); and (d) HRM policies and practices (staffing, skills).

Ultimately, whether process indicators or output indicators are most useful will depend upon the objectives of the user. Where the information is required for diagnostic or dialogue purposes, indicators which measure particular reform processes are likely to be most helpful because they will help the user to pin-point exactly what is wrong and in need of improvement. To do this, the indicator must assess the processes in-use as well as those defined in-law or regulations. An example is provided in Box 4.

BOX 4 – PARTICIPATION OF WOMEN IN THE CIVIL SERVICE: EXAMPLE OF INDICATORS

Public Administration Reform objective
To ensure the representation of women in senior management broadly reflects their composition in society

Output indicator
% of vacant management positions filled by women

Process indicators
- New selection criteria which recognise women's contribution
- Flexible employment policies which allow women more freedom to choose their hours of work
- Equal training opportunities which permit women with child-rearing responsibilities to attend training close to home
- External competition for jobs to widen the pool of potential female applicants

---

Process indicators are therefore the "actionable governance indicators" that users have been demanding. In contrast, output indicators are far more useful for monitoring and accountability purposes because they tell the user what results the reforms are achieving without providing unnecessary detail on specific processes. That level of detail can be provided subsequently if it is evident that the intended results have not been achieved.

1.12 Combining quantitative data with qualitative data

Generally, PAP assessment tools draw from five sources of information: written surveys (e.g. self-assessments), dialogue-based processes (e.g. interviews, focus groups), indirect observation (e.g. what the assessment team sees when it is in a government building), direct observation (e.g. shadowing a staff member over a day to see how business processes work), and quantitative data (e.g. on services provided such as patients served at a health post).

It is often tempting to believe that tools and sources which generate metrics (numbers) are somehow providing more objective data. Unfortunately, such objectivity is largely illusory. Objectivity will depend on whose opinion is being quantified. For example, survey data, while presented in a quantitative fashion, is perception-based. Also, many qualitative assessments (such as PEFA) are converted into scores or ratings, usually based on clear scoring criteria. Objectivity will also depend upon the accuracy of the source when data is presented in the form of raw statistics. A good example is the World Bank's wage-employment database where it is always problematic for country economists to obtain accurate and up to date employment data from government HR information systems. In addition, cross-country comparisons are risky because of the different national definitions of civil service employment used. It is essential therefore for the user to dig beneath the surface to determine how the numbers are generated.

Quantitative data is not necessarily more accurate than qualitative data

Most of the tools and sources rely heavily upon qualitative data which is essential for measuring change processes in each of the thematic areas. For some tools, qualitative data is reported in a narrative format and, for others, it is converted to a score. Both qualitative and quantitative information are necessary for measuring PAP components, but neither is necessarily superior to the other. There is often less risk attached to qualitative data, whereas quantitative data can be spurious or inaccurate depending upon the source. Ideally, qualitative data should be substantiated by quantitative data. In practice, it is often difficult to generate the quantitative data required during the short time in which the typical assessment takes place, unless government's information systems have already been designed to produce this information.

Perceptions of key stakeholders represent valid data for PAP assessments

One type of qualitative data – perceptions – is often treated dismissively, especially by professionals who are trained to search for numbers. Yet, the opinions of customers (on services delivered) and the views of civil servants (on human resource management practices) provide valid data concerning the strengths and weaknesses of particular PA themes. Because the views of stakeholders affected by PA performance are extremely relevant, such soft measures can be legitimately regarded as hard data.

1.13 What are the most appropriate techniques to collect information?

Methodology refers to the techniques and processes used for the collection of data, its subsequent analysis and the provision of reports for the user.

As summarised in Table 3, a wide range of methods are employed by the different information sources and assessment tools. The most common methods include surveys, questionnaires, interviews, desk research and discussion. Many assessment tools use a combination of methods. For instance, the institutional assessment of the civil service system employs an assessment questionnaire, desk research, interviews and expert panels. In contrast, the information sources rely exclusively on either desk research or a combination of surveys and desk research. A few assessment tools are designed as self-assessment instruments and others incorporate a self-assessment option.

One advantage of using multiple methods is that they have the potential to generate a wider range of information from a number of different sources, and they permit the validation and triangulation of data. However, employing many methods can be costly and complex to administer, and in the end quality may be compromised. Also, a government should not attempt too much research if it lacks the capacity to process the data into meaningful information and provide feedback to those who volunteered the information. For instance, carrying out extensive surveys of civil servants and citizens may sound like a laudable endeavour, but it can backfire if the respondents do not see the results. It will simply confirm their prejudices that the government was not really interested in acting on the data in the first place.

Some methods can be used to complement each other, rather than to generate stand alone data. For example, the Common Assessment Framework uses a self-assessment to generate preliminary data which can be used as a basis for discussion. Such an approach enriches the analysis of problems and helps to build a consensus on the top problems that need to be tackled. This is a "light" method which can generate valid information from respondents who are familiar with the issues, without the need for extensive research to assemble quantitative evidence.

Whilst surveys and questionnaires have the potential to obtain the views of many respondents in a relatively short period of time, it should be borne in mind that these will inevitably be carried out on a sample basis. As far as possible, the sample should be selected at random. Selection biases can occur if the respondents are senior civil servants who are less likely to be critical of its policies and processes. It will also be necessary to conduct statistical analyses (of the standard error) to be confident that the results of the sample reflect those which prevail in the total population.

However, if the user does not have the time, the skills or the resources to design and administer a fully-fledged statistical survey, one option is to carry out a rapid review to produce a snapshot of the key issues. Such a review would involve semi-structured interviews with a few key informants which aim to cover the full scope of the assessment. The rapid review can serve as a funnelling technique which enables critical issues to be investigated further using other methods, such as interviews or focus groups.

Self-assessments are worth considering as a contribution to a discussion, rather than a final verdict

Self-assessments can work well where there are explicit criteria to enable respondents to form an opinion. Even so, there may be a tendency for respondents to inflate the rating if they believe they will be held accountable for poor performance. Also, the majority of people tend not to be very self-critical. These problems can be countered to some extent by including different perspectives – for instance, civil servants who are affected by but not directly responsible for the thematic area – and insisting that individual assessments are justified. Our view is that self-assessment data is potentially useful because those generating the data have a real stake in the outcome, with the caveat that it is clearly understood that the information will be used for improvement, and not accountability, purposes.

The final column in Table 3 shows who provides data for the assessment tools and information sources. The main respondents are: (a) independent experts and consultants; (b) agency staff; (c) government officials; and (d) external stakeholders.

Involving external and internal consultants in a team approach can enrich the quality of the assessment

Many of the assessment tools rely to some extent on the judgments of independent consultants, experts and academics.[22] There are two main reasons for this. Firstly, they are deemed to have expertise in the theme being assessed. Secondly, their opinions are perceived as independent, hence more objective, than those of the responsible government officials. Using 'outsiders' may produce an accurate assessment, but unfortunately this may not result in action by government if external experts are not seen as credible and trusted sources. For this reason, a national or regional expert who has a good understanding of the local context is often preferable to an expert who is external to the region. For example, a PAP assessment in an EU candidate country to assess compliance with EU accession criteria might be

---

effectively led by an expert from a new member state of the region, but less so by a national expert, who might not be favoured in view of an affiliation with previous governments. There is also a risk that national consultants might be subjected to undue pressure by government officials seeking a favourable assessment. A pragmatic solution is to embrace both foreign (ideally regional) and local consultants working as a team.

Government officials have a critical role to play in the assessment

Government officials tend to be used far less frequently as data sources than independent experts and donor agency staff. In fact they are not used at all in any of the information sources. From one viewpoint, this is perhaps surprising because government officials are usually far more knowledgeable about their problems than a foreigner who jets in for a few weeks. Indeed, an external consultant will usually be obliged to obtain much of the information for his/her assessment by talking to local government officials! It seems that the agencies who have developed these tools either believe that government officials have insufficient expertise or that they do not have enough time to devote to a comprehensive assessment. Fortunately, this view is now changing with new tools being designed for the participation of government officials.

Joint assessments carried out jointly by external consultants and government officials have the greatest potential to generate balanced assessments[23]

Joint assessments have two main advantages. Firstly, they are likely to generate high quality data and secondly, they increase the likelihood that the results will be acceptable to government. However, whether a joint assessment can be successfully conducted will depend upon the skills available within government. In practice, it is often difficult to find enough civil servants who are capable and willing to dedicate themselves full-time to the task. And some managers are often reluctant to release the most qualified staff because they are in high demand for operational tasks. Another, less significant, issue is that some civil servants may be reluctant to be seen to be judging their peers or superiors. Where these practical problems can be overcome, joint assessments work best if they are government led, with external consultants used in an advisory capacity. In this context, the Government of Zambia recently led a PEFA assessment, although the majority of such assessments have been consultant-led.

360 degree feedback is a potentially useful approach, but the cultural limitations need to be recognised

Is it possible to the joint assessment approach a step further and adopt a 360 degree feedback instrument which takes account of the views of all stakeholders? Taking a particular PAP theme as the target for assessment, the 360 degree approach (Figure 3) would collect the views of ministers, experts (local or foreign), civil servants and citizens (directly or through civil society organizations.) None of the assessment tools take this panoramic perspective, although the Asian Development Bank's Country Governance Assessment comes close. The choice of 360 degree feedback is especially appropriate for assessing the Government policy making process and it is also a relevant perspective for civil service management. The approach is less suited to the assessment of public financial management where the internal processes under review may not be visible to either civil servants or citizens. 360 degree feedback is therefore a useful perspective depending on the knowledge and experience of the respective stakeholders.

However, hierarchic and autocratic governments which have limited tolerance for criticism may be reluctant to accept the views of citizens and civil servants as valid or legitimate. And in many countries, civil servants and citizens may be reluctant to voice their true opinions where they fear that the questionnaire or survey is not truly confidential. This can even occur where the necessary safeguards have been put in place to assure the anonymity of opinions expressed. Such governments may also reject the views of experts. In such circumstances, where governments are generally reluctant to listen to "outsiders" because they fear criticism, a self-assessment represents a practical entry point to an assessment. The assessment could constitute the basis for a discussion facilitated by a trusted and credible "outsider."

---

[23] Government officials may include those with direct responsibility for the thematic area and other civil servants with good diagnostic skills

Figure 3 – 360 degree feedback
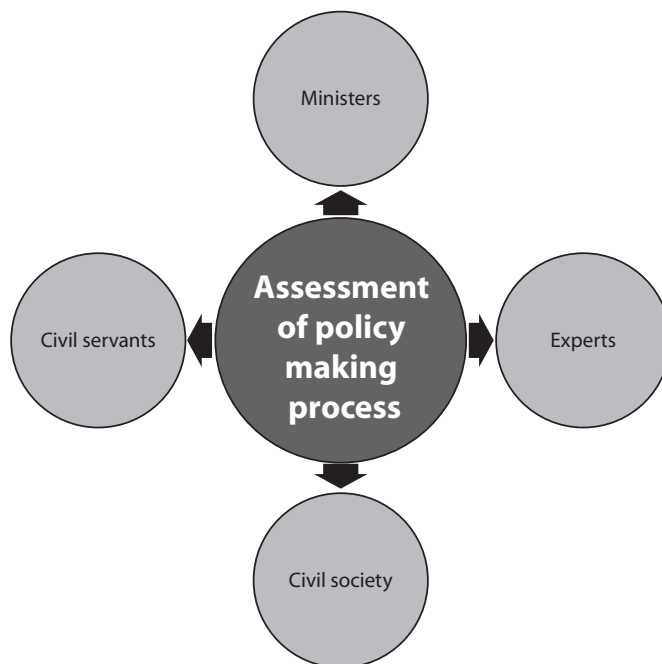
It doesn't really matter who produces the report, so long as the assessment process agreed by stakeholders at the outset is followed

Most of the assessment reports and information sources are produced by the international agencies who own the tools. This leaves open the possibility that a recipient government will reject the findings because it is seen to represent the agency's view. This is more likely to happen if there is evidence that the agency has manipulated some of the findings after the completion of the assessment. However, provided the conclusions in the report are clearly based on evidence which was discussed and agreed with the host government, then it should not matter who formally authors the final report. In fact, those government officials who participated in the assessment may actually prefer the report to be produced by an international agency. The agency is in a safer position to present unpalatable truths to government, which officials will sometimes be reluctant to communicate themselves.

1.14 Using composite indices for cross-country comparisons: Uses and pitfalls

Information sources and many assessment tools use ratings both for benchmarking against a standard over time and for cross-country comparisons. This does not pose a problem where the rating simply converts an assessment on a particular indicator to a score. But where the ratings on indicators are aggregated to give a composite score, then the practice is highly questionable because it obscures what is being measured. If the rating is to be linked to some form of recognition – such as the UK's Self-Assessment Tool for Customer Service Excellence – then scoring is both necessary and justifiable. However, where the assessment is to be used for the purpose of diagnosis and dialogue, then scoring is less relevant.

Ratings are superficially attractive, but there are inherent risks

Scoring can help the user to compare performance across a number of different dimensions. But it is critical to understand the basis of the scoring. For instance, PEFA provides detailed guidance to help assessors differentiate between an A, B, C or D for each indicator, thereby requiring the ratings to be evidence-based. The process is therefore transparent and allows government to challenge an assessment based upon an examination of the evidence. In contrast, whilst the UK Civil Service's Capability Reviews uses a common set of behavioural criteria for the departmental reviews, the methodology gives little concrete guidance for differentiating between the five assessment categories – strong, well placed, development area, urgent development area, serious concerns. In its defence, the Capabilities Review Team in the UK Civil Service moderates the exercise for the civil service as a whole to ensure that the judgements of the departmental review teams are broadly consistent. The central oversight function is of course not an option for a user that wishes to carry out such as review for a single department.

A major disadvantage of the use of ratings is that it can divert the energies of the parties involved in the assessment. Too much time is spent negotiating and deciding upon the indicator score and consequently not enough emphasis is placed the nature of the weakness and opportunities for improvement. Where ratings are provided, there is strong case for supporting these with a narrative assessment, as is provided in the PEFA assessment tool.

1.15 How can ownership be fostered?

Users should participate in the design of the assessment process

In the foregoing three sections guidance has been provided to enable users to: (a) select tools which meet their objectives; (b) choose relevant indicators; and (c) design a methodology for collecting data on these indicators. Ownership of the assessment process is therefore achieved when the user is able to contribute to these "technical" decisions. People usually commit to what they help to create. An international agency may believe it has the right answer on these questions for its government client. But if this view is imposed, or perceived to be imposed, then compliance rather than ownership will be the inevitable result.

The quantitative service delivery surveys, the public expenditure tracking survey and citizen report cards insist upon user participation because they must be customised to the specific public service that is being assessed. And whilst those tools which measure the core PAP themes appear to have fixed indicator sets, the user is encouraged to propose including additional indicators and removing existing ones.

Participation is essential because tools will inevitably need to be customised

A government may wish to use an assessment tool as part of a diagnostic exercise which feeds into a PAR strategy development process. The measures and methods employed in the tool should be designed to fit the specific objectives of the assessment, not the reverse. There is an even stronger case for government to be involved in the design of an assessment tool which is intended for monitoring an ongoing reform programme. Here it is essential that the indicators are aligned with the government's reform strategy. The measures should be derived directly from the specific programmes and targets previously agreed upon in the strategy.

Ideally, users should also contribute to the actual assessment

User engagement in the design process is the best insurance for generating results that are accepted and acted upon. However, the findings may still be rejected where the user decided to contract out the entire assessment to a third party. This is the classic "turnkey" operation. This problem is best avoided by involving government officials in the collection and analysis of data, or better still in contributing to the actual assessment. People rarely argue with their own data. Valid and actionable assessments cannot be constructed on the basis of the opinions of experts alone.

In the long-term, Governments must build their capability to conduct assessments on a regular basis

Most tools are designed as one-off tools to be applied at discrete intervals, resulting in ad hoc data collection exercises. Whilst it will always be necessary to conduct special surveys, Governments must also strive to build their own information systems which can routinely generate key data on PAP. Surveys can be prohibitively expensive to administer, especially those of citizens and civil servants. However, surveys of citizens do not need to be undertaken too frequently and the costs of surveying civil servants can be minimised by designing questionnaires which can be completed on-line.

It will take time for governments to build their capabilities to conduct PAP assessments without external assistance. But they can learn quickly by insisting that government officials participate in joint assessment teams working alongside more experienced consultants.

Table 3 – PAP Assessment Tools and Information Sources: A Snapshot

Ref	Tool & producer	Themes	Users	Objectives[24]	Methods	Respondents[25]
			Assessment tools			
1	**Quantitative Service Delivery Surveys,** by World Bank	• **Service delivery** *Also useful to inform public expenditure management reforms*	• Government • Service providers • Donors	• Diagnosis • **Monitoring**	• Provider records • Interviews • Surveys	• **World Bank staff** • Provider staff • Clients
2	**Public Expenditure Tracking Survey,** by World Bank	• **Public financial mgt** *To examine efficiency of expenditure and "diversion" of resources*	• Donors • Government	• **Monitoring** • Diagnosis	• Interviews • Provider records • Surveys	• **World Bank staff** • Provider staff • Clients
3	**Citizen report cards,** by Public Affairs Centre	• **Service delivery** *For beneficiaries to engage in dialogue with service providers*	• Civil society • Government • Service providers • Service beneficiaries	• **Dialogue and joint decision making** • Benchmarking over time • Monitoring • Diagnosis	• Interviews • Focus groups	• **Civil society** • Clients
4	**Country Governance Assessment,** by Asian Development Bank (ADB)	• **Policy-making** • **CSR** • **PFM** *Also assesses local governance*	• Donors • Government	• **Diagnosis** • Dialogue and joint decision making • Resource allocation	• Desk research • Field visits • Interviews • Workshops	• Donor staff • Consultants • Government • Local stakeholders
5	**Common Assessment Framework,** by European Institute of Public Administration (EIPA)	• **Leadership** • **CSR** • **PFM** • **Policy-making** • **Service delivery** *For use in all parts of public sector, at national, regional and local levels*	• Public organization • Government	• **Diagnosis** • Cross-comparison • Benchmarking	• Self-assessment questionnaire • Discussion (or facilitated discussion)	• **Public organization** • Managers & staff
6	**Capability reviews,** by UK Civil Service	• **Leadership** *To assess leadership, strategy and delivery*	• Public organization • Government • Clients/Service beneficiaries	• **Diagnosis** • Monitoring	• Interviews	• **External experts** • Managers
7	**Self-Assessment Tool for Customer Service Excellence,** by UK Civil Service	• **Service delivery** *To assess customer focus and customer engagement*	• Public organization	• Monitoring • **Diagnosis** • Benchmarking over time • Cross-comparison	• Self-assessment questionnaire • Provider records	• **Provider staff**
8	**Public Expenditure and Financial Accountability (PEFA),** Multi-institution project (World Bank, IMF, EC, DFID, SPA)	• **PFM** *To assess credibility, comprehensiveness and transparency of budget*	• Donors • Government	• **Monitoring** • Benchmarking over time • **Dialogue and joint decision making**	• Self-assessment questionnaire • Interviews • Desk research	• **Donors** • **External consultants** • **Government officials**
9	**Public Officials' Survey, by World Bank** – Netherlands Partnership Program (BNPP)	• **CSR** • **PFM** *To assess the incentive structure in the public sector*	• Government	• **Diagnosis**	• Surveys	• **Public officials**
10	**CONTACT,** by UNDP	• **PFM** *To assess financial accountability and integrity systems*	• Government	• **Diagnosis**	• Self-assessment questionnaire	• **Government officials**
11	**Evaluation Matrix of Civil Service Human Resource Management in the EU,** by World Bank	• **CSR** • **Policy-making** *To help EU candidates specify administrative development targets necessary for EU accession and global competitiveness*	• Government • Donors	• **Diagnosis** • Monitoring	• Assessment questionnaire • Discussion	• **World Bank staff [EC too]** • **Government officials**
12	**Control & Management Systems Baselines,** by SIGMA (OECD & EU)	• **CSR** • **PFM** • **Policy-making** *Used by EC for assessing EU candidate countries*	• European Commission	• Diagnosis • **Monitoring**	• Desk research • Interviews	• **SIGMA staff** • Government officials • Donors
13	**Human Resources Self-Assessment Guide,** by State of Texas	• **CSR** *For self-assessment of HRM by state agencies and universities*	• Public organizations • Universities	• Diagnosis • **Monitoring**	• Self-assessment questionnaire • Desk research	• **Managers & staff**
14	**Human Resource Management (HRM) Assessment Instrument,** by Management Sciences for Health	• **CSR** *For self-assessment of HRM by NGOs and public sector health institutions*	• Public organizations • NGOs	• Diagnosis • **Monitoring**	• Self-assessment questionnaire • Desk research • Discussion	• Managers & staff

Ref	Tool & producer	Themes	Users	Objectives[24]	Methods	Respondents[25]
15	**Analytical Framework for Institutional Assessment of Civil Service Systems,** by Inter-American Development Bank (IADB)	**• CSR** *To assess civil service systems in countries located with IADB's scope of action*	• Researchers • Donors	**• Diagnosis** **• Cross-comparison**	• Assessment questionnaire • Desk research • Interviews • Expert panels	**• External experts** • Agency staff • Public officials
16	**Engendering Budgets: A Practitioners' Guide to Understanding and Implementing Gender-Responsive Budgets,** by Commonwealth Secretariat	**• PFM** *To assess gender-responsiveness of budgets*	• Government • Civil society	**• Monitoring**	• Surveys • Desk research • National budget • National statistics	**• Government officials** • Households **• NGOs**
17	**Open Budget Initiative,** by International Budget Project	**• PFM** *To assess a government's commitment to budget transparency and accountability*	• Citizens • Civil society • Legislators	**• Dialogue and joint decision making** • Cross-comparison • Diagnosis	• Desk research	• Local experts
18	**HRM Actionable Governance Indicators,** by World Bank	**• CSR** *To assess reforms 'in law' and 'in practice'*	• Government • Donors	**• Diagnosis** • Monitoring	• Desk research	• Donor staff • Government officials
19	**Diagnostic Framework for Revenue Administration,** by World Bank	**• PFM** *To provide a comprehensive diagnostic framework for revenue administration*	• Government	**• Diagnosis**	• Self-assessment questionnaire	**• Government officials**
			Information sources			
20	**Government at a Glance,** by OECD	**• CSR** **• Policy-making** **• PFM** **• Service delivery** *Covers OECD member countries, no overall score and no rankings*	• Government • Practitioners	**• Monitoring** • Dialogue • Cross-comparison	• Survey • Desk research	• OECD staff • OECD member States
21	**Management Index, Bertelsmann Reform Index,** by Bertelsmann Siftung	**• CSR** **• Leadership** *Assesses 'executive capacity' and 'executive accountability' in 30 free market*	• Donors • Government	• Benchmarking • Monitoring **• Diagnosis**	• Survey • Desk research	• In-country and external experts
22	**Cross-National Data on Government Employment & Wages,** by World Bank	**• CSR** *Database on public sector employment and pay in more than 200 countries*	• Government • Donors • Researchers	• Cross-comparison **• Monitoring**	• Desk research	• World Bank staff
23	**World Competitiveness Yearbook, Government Efficiency,** by IMD	**• PFM** **• Policy-making** *Assesses contextual factors impacting competitiveness of enterprises in 55 countries*	• Government • Investors • Donors	**• Cross-comparison** • Monitoring	• Survey • Desk research	• Leading scholars • IMD staff
24	**Governance Matters, Government Effectiveness,** by World Bank	**• CSR** **• Policy-making** **• Service delivery** *Worldwide coverage; aggregates 35 separate data sources measuring perceptions of the quality of public service*	• Government • Donors • Researchers	**• Cross-comparison** **• Monitoring**	• Surveys	• Firms • Businesspeople • Experts • Households
25	**Country Policy and Institutional Assessment (CPIA)** Criteria 13: Quality of Budgetary & Financial Management Criteria 15: Quality of Public Administration, by World Bank	**• PFM** **• CSR** **• Policy-making** **• Service delivery** *To assess quality of policies and institutional arrangements in IDA-eligible countries (the key elements that are within a country's control*	• Government • Donors • Researchers	**• Cross-comparison** **• Monitoring**	• Questionnaire	• World Bank country teams

Table 3 – PAP Assessment Tools and Information Sources: A Snapshot (continued)

Ref	Tool & producer	Themes	Users	Objectives[24]	Methods	Respondents[25]
26	**Integrity Indicators,** by Global Integrity	• **PFM** • **Leadership** • **Policy-making** • **CSR** *To assess the existence, effectiveness and citizen access to national anti-corruption mechanisms*	• Government • Donors • Researchers	• **Diagnosis** • Benchmarking • Cross-comparison	• Questionnaire	• National experts • National journalists

[24]   Primary objective is in bold text
[25]   Those responsible for producing the assessment report shown in bold

A Users' Guide to Measuring **Public Administration Performance**

CHAPTER 2
GOOD PRACTICES: GUIDANCE FOR GOVERNMENTS

The guidance is aimed at governments specifically and at stakeholders working in partnership with governments.

Getting started

- Be clear about your objective first before you choose an existing tool or attempt to develop your own
- If you are looking for an assessment tool for diagnostic or dialogue purposes, consider an existing tool which matches your chosen theme as a starting point
- If you want to monitor an ongoing reform programme, develop your own measurement tool. Clarify your vision, develop your strategy map, and construct your indicators to reflect the key drivers
- Don't restrict your choice to a single tool, as they are not mutually exclusive. Where appropriate combine the good features of different tools
- If you want to develop a customised tool, don't start with a blank sheet of paper. Start by reviewing the indicators in an existing tool, retain them or discard them as appropriate and include some locally relevant indicators of your own

Measurement focus

- In general, focus on measuring change processes and their intermediate outputs, avoiding indicators too far down the service delivery chain. For change processes, identify the key criteria. Don't attempt to measure everything
- Measure only what you can change when you are conducting a diagnostic, so identify actionable indicators. But avoid getting drawn too deeply into assessing processes which are only loosely connected to the outcomes you are seeking. Don't lose sight of the big picture
- Don't be afraid of building your values and beliefs into your indicator set, because measurement should not be a neutral technical exercise. Make sure though you are explicit about your normative standpoint, otherwise you may be accused of having a hidden agenda
- When assessing PA performance, qualitative data tends to be more useful and accessible than quantitative data. Use quantitative data when available to support qualitative data, but don't be fooled by misleading metrics
- Whilst your policies to improve PA performance will remain constant, the pathway may change from year to year. Adjust your indicators accordingly to reflect changes in your operational priorities
- Acknowledge the level of maturity of your reform programme. Customise the assessment tool to make sure it is measuring your current reform priorities, not those that you may want to address in the distant future
- Make sure that your indicator set is expressed in a language that you and your respondents can understand. Don't blindly copy the terminology used by the tool's authors
- Choose the right time to do the assessment to maximise the chances that the data will be acted upon. It is usually good practice to align measurement with the government's budget process and calendar

Methodology

- Start with the data you have, not the data you would like to have in an ideal world
- Use surveys and questionnaires to obtain views from many stakeholders in a relatively short period of time. But recognise that surveys of citizens and civil servants may not generate valid or accurate data because: (a) respondents may be concerned about anonymity and confidentiality; (b) in many cultures people are reluctant to be critical of their superiors

- Follow up surveys with focus groups to interpret findings and to begin searching for solutions
- Recognise that traditional surveys can be expensive to design, administer and process. Consider inexpensive rapid reviews instead, but validate the findings using other methods
- Use self-assessments as an input to discussion, not as a final verdict
- Avoid placing too much emphasis on scoring and ratings, as arguments over the score can distract attention from what you need to do to improve (unless you are only interested in benchmarking)

Data sources

- Involve civil servants in assessing their own systems, ideally junior or mid-career, not too senior, without a vested interested in the status quo
- Embrace joint assessments carried out by government officials and consultants (or agency staff.) Employ both external and internal consultants working as a team to obtain best value from independent assessors
- Don't ask civil servants or citizens for their opinions if you do not intend to give them feedback on what they said. Otherwise they may not volunteer this information again!
- Consider 360 degree feedback to provide a balanced picture from a range of perspectives. But don't impose it where respondents have no knowledge or legitimate interest in the theme to be assessed

Ownership and sustainability

- Government ownership of the assessment tool is essential when its purpose is to monitor PA performance or an ongoing reform programme, because the indicators must then be aligned with government's reform strategy
- Involvement of government in the design of the assessment tool – indicators and methodology – is most critical factor in creating ownership. Ownership of the actual data is important too, but it may be less critical if there has been involvement in the design
- Graduate towards improving information systems and capacity, and away from one-off assessments. The problem with external assessments is that the data may not be available in particular countries

CHAPTER 3
CASE STUDIES

The four case studies described below have been designed to illustrate how some of the good practice guidelines can be applied in particular contexts. All of them are fictional, but they nonetheless draw upon actual experiences in designing and conducting assessments.

Case study 1: 360 degree feedback tool for Government policy making

Lamin is a university professor with a PhD in political science who has been an academic for his entire career. He has been engaged by the Head of the Civil Service to design a diagnostic tool to assess the strengths and weaknesses of the Government's policy making process. If the client likes the product, he expects to be asked to lead the diagnostic process itself. The Head of Service, as Secretary to Cabinet, is coming under increasing pressure from the President because of the sheer volume and poor quality of Cabinet memoranda that are being submitted on a weekly basis. As a result the Cabinet is unable to conduct its business effectively. Decisions are been rushed because of the lack of time and some papers are being returned to the originating ministry for further clarification prior to their re-submission.

Lamin meets the Head of Service to make a start on this assignment. He is acutely worried about the President's expectations. Is he looking for a quick fix solution that will give his Cabinet some urgently needed breathing space? Or does he want a more durable and sustainable solution? The Head of Service confirms that the President wants a quick solution but, at the same time, because he was only elected to office last year, he wants a solution that will work for the whole of his term. In the course of the conversation, the Head of Service reminds Lamin that the President was elected on a populist platform, winning many votes in the rural areas. There were many promises in the manifesto of the ruling party to improve basic education and health services.

Fortunately, Lamin has recently read UNDP's Users' Guide for Measuring PAP and he knows that there are some existing tools which can be used to measure the Government's policy making process. More important, he recalls the interesting concept of 360 degree feedback because this provides a unique opportunity for the ordinary citizen to express an opinion on government's reform processes. He believes it will be a good idea to use this technique in this assignment. He decides therefore that government should obtain the views of four different sources on the policy making process – ministers, citizens, civil servants and outside experts. But will this be sufficient? He then remembers that there is a Cabinet Secretariat which is responsible for screening Cabinet memoranda originating from ministries before they are placed on Cabinet's agenda. He is sure they will have a view on what is wrong with the process and he is surprised that the Head of Service did not mention this at their meeting. He phones the Head of Service and asks to meet the Cabinet Secretariat. He is told that this would be a waste of time because the staffs are all very senior but useless bureaucrats. Oh dear! Lamin goes back to the drawing board.

He feels he must disagree with his friend the Head of Service, who just wants to impose a solution upon an unwilling Secretariat. Perhaps he believes that the proposals will carry more weight if they are made by an eminent university professor! He is not so sure about this. Although he has never been a civil servant, he knows that civil servants often have tremendous power to block change, even when politically sanctioned. Even if, as it seems, the Secretariat staff are a major part of the problem, he believes they must also be a part of the solution. First, he considers arming himself with the SIGMA checklist and interviewing members of the Secretariat. But he has second thoughts because he fears that the Secretariat will be defensive. He decides instead to adopt a self-assessment process which will give them an opportunity to present their side of the story.

He persuades the Head of Service to allow him to go ahead with the self-assessment, as a kind of pre-diagnostic research. He believes it is essential to talk with the department responsible for the process before designing the 360 degree feedback tool. He is a little nervous but, to his surprise, the Secretariat welcomes his visit and his open approach to questioning. It turns out that they themselves are feeling overburdened with the volume of work, having to deal with last minute requests from Ministers to include memoranda in Cabinet's agenda and having little time to improve upon the poor quality of Cabinet papers which are submitted by ministries. Although he feels that to some extent the Secretariat may be trying to pass the blame to others, it is clear that the causes of the problem are complex and that there are many actors (ministers, civil servants in ministries) who are a part of the problem. He now feels vindicated in his decision to use a 360 degree feedback tool and he is easily able to persuade the Secretariat that this is the right way forward.

Lamin's next challenge is to obtain the support of the Secretariat in designing the four instruments for ministers, civil servants, civil society and external experts. This is important because they have far better knowledge of the existing policy making process than he does. This time he faces no obstacles, since by now, the Secretariat staff realise he is not looking for a scapegoat. Together, they quickly decide that semi-structured interviews, structured interviews and focus groups are the most appropriate methods for ministers, civil servants and citizens respectively. To obtain the views of external experts, they decide to ask one local consultant and foreign consultant who has previously worked in the country to carry out an assessment based on the SIGMA standards. But they stress that the SIGMA standards should be used as a guide not a template, and that the experts will have discretion to include other parameters that they consider relevant.

Finally, individual members of the Secretariat are tasked to work with Lamin on the three instruments for ministers, civil servants and citizens. The Secretariat is keen to be involved in the diagnostic itself and it is decided that individual staff will conduct the structured interviews for civil servants in line ministries who have participated in the policy development process. Lamin, however, will take responsibility for interviewing ministers because they will in all likelihood refuse to be interviewed by a civil servant! The final issue to be resolved is who will conduct the focus groups for citizens. They decide that civil servants neither have the skills nor are they likely to be trusted by citizens most of whom will have had very bad experience with public services. Instead, they opt to invite a local NGO to conduct the focus groups, who will work with Lamin and the Secretariat team to agree the issues that need to be raised. Participation in the focus groups will be organized by the NGO and will be kept absolutely confidential. Criteria for participation will be agreed with Lamin to ensure a cross-section of citizens in different poverty bands.

Lamin and the Cabinet Secretariat then prepare a joint plan and budget for administering the 360 degree feedback instrument which is readily accepted by the Head of Service and Cabinet. The results of the assessment conclude that: (a) citizens have few opportunities to articulate their opinions on draft policy proposals which are emerging; (b) many proposals submitted to Cabinet are management and administrative matters which should be disposed of by Permanent Secretaries; and (c) ministries lack the capacity to analyse and develop high quality policy proposals. The findings lead to the development of an improved policy management process, which is documented in an updated Cabinet handbook, and the conduct of a policy management training programme for senior managers and planning staff in ministries.

Case study 2: Designing an Employee Attitude Survey

Charles is an international consultant who specialises in human resource management, with substantial experience of civil service management programmes in developing countries. In response to a request by government, he is contracted by UNDP to design and administer a survey of civil servants to identify their views on government's human resource management policies and processes. UNDP is aware that the only off-the-shelf surveys available are proprietary tools owned by consulting companies for which a substantial fee would be charged. Charles has conducted similar surveys in other countries and is aware of the challenges and the pitfalls.

The Director of the Central Personnel Department is in a hurry to conduct the survey because he is experiencing pressure from the civil service association to take into account the views of civil servants in the development of a new public administration reform strategy. The easy approach would be to use a questionnaire used in another country which is the preference of the Director of Personnel. Charles explains that, whilst there are standardised attitude surveys that have been applied in different organizational contexts, these rarely translate across national borders. It is clear though that the time and budget will not allow the development of a questionnaire from first principles. As a compromise, it is agreed that a questionnaire which Charles has used in another country will be used as the basis for developing the questionnaire. The Director of Personnel decides that the questionnaire should be tested on the two most senior personnel specialists in his

department, since they are the experts in such matters. No-one understands the rule book better than they do! Charles has a slightly different point of view. He argues that personnel specialists in line ministries also have a contribution to make because they understand better how the current practice affect civil servants. The Director of Personnel agrees that they should include two personnel specialists from line ministries in the group.

First, however, Charles spends time with the Director of Personnel to gain a better understanding of the existing human resource management practices. He also reviews the Civil Service law, various HR policy statements in force, and the civil service regulations. He is now better prepared to have a discussion with the group of personnel specialists. The group discussion goes well and a number of changes are made to the prototype questionnaire. Some new questions are introduced, others are removed and the terminology is given a significant overhaul so that it is consistent with the language used in the personnel regulations with which civil servants are familiar.

The group then discusses how the questionnaire should be administered and how the sampling should be done. Everyone agrees that the respondents should be selected at random but at the same time should be representative of all salary grades. There are different viewpoints about the administration of the survey; some believe it should be distributed by the personnel department in each ministry with an instruction to the respondent to return it at an agreed future date; others consider that that the respondents should be gathered in a meeting room and complete the questionnaire before they leave. Charles explains that in his experience the second approach works best because the personnel department has an opportunity to explain the purpose of the questionnaire and to assuage any particular concerns about confidentiality on the spot. It also guarantees a 100% response rate! In the end everyone agrees to this approach.

Charles then suggests that the questionnaire should be pre-tested on a few civil servants to confirm that the questions are easily understood. One of the personnel specialists proposes that the Chairman of the Civil Service Association ratify the questionnaire before it is pre-tested. Everyone accepts that this is a good idea. At this juncture, Charles proposes that the questionnaire be combined with a discussion group where some respondents interpret some of the findings and can express their view more fully. In this way, the quantitative data from the survey can be supplemented with qualitative data from the follow-up discussions. It is agreed that the discussion group should include volunteers only and that they should be given the summary findings of survey beforehand.

Now that the group has agreed the way forward, they present their recommendation to the Director of Personnel who promptly endorses it. Finally, they go to meet the Chairman of the Civil Service Association to explain the overall approach to the survey, to show him the questionnaire and to obtain his support for the exercise.

The results of the questionnaire survey, which are reinforced by the discussion group, reveal widespread dissatisfaction with the access of civil servants to training opportunities, promotion decisions and the operation of the current performance appraisal system. The Director of Personnel uses the evidence of the survey to design an improved performance management system for the civil service linked to development and to develop a new promotions policy emphasizing merit as opposed to seniority as a criterion for promotion. The employee survey is repeated two years later and reveals improved levels of satisfaction among civil servants.

Case study 3: Designing a Citizen Report Card

Sarah is the Chairperson of the management committee for a government health clinic. The main role of the management committee is to oversee the collection of user charges and to decide how this income should be used to improve the facilities provided at the clinic. Sarah was elected to this position at a meeting convened by the local chief three months ago. Some of Sarah's neighbours have come to see Sarah about the deteriorating standards of care at the clinic. The clinical staffs are often absent, they can be disrespectful to patients and waiting times are long. There are rumours that one old woman passed away on a bench waiting to see a doctor. They believe she had a stroke.

Sarah decides that she must see the medical officer in charge of the clinic to bring his attention to the concerns that have been expressed by the local community. He is reluctant to meet her, claiming that he is too busy, but in the end he agrees to meet at the end of the day when the number of patients has dwindled. The doctor likes Sarah because she is an educated woman, unlike most of the other people in leave in the nearby villages. He tells Sarah that he and his staff are overloaded and work long hours for which they are not rewarded. He admits that he has a problem with staff absenteeism, but is powerless to do anything about it. He mentions that one of his enrolled nurses is looking after a sick relative who is very

demanding. He is dismissive of the assertion about the poor standards of care, saying that most the local people are ignorant. Finally, he tells Sarah politely that the standards of care are no business of the management committee. The committee simply exists to decide how the user income should be spent.

Sarah comes away deflated from her meeting with the Medical Officer in charge. She decides to convene a meeting of the management committee to see whether any of her colleagues have any ideas about what can be done. Fortunately, the Treasurer has a relative who works at the health planning unit in the Ministry of Health in the capital and suggests that they go and meet her. Sarah and the Treasurer travel to the capital and meet the planning officer who tells them about a new initiative in the Ministry to introduce Service Charters for all health facilities. A Service Charter will state the services provided at the facility, it will specify service standards (for instance, how long a patient can expect to wait to see a doctor of a nurse), and how staff should behave when dealing with patients. The intention is that the Charter will be displayed as a poster at the clinic so that all patients can see it.

Sarah thinks the Service Charter is a great idea which can help to strengthen the accountability of health personnel to the community. Unfortunately, the Ministry does not plan to introduce the Service Charter for at least another 6 months. Sarah and her community cannot wait that long, so the planning officer advises them to speak to a young lecturer at the university she knows who did his PhD research in India on the use of citizen report cards. They meet the lecturer who agrees to help them with the design of a survey at no charge because he wants to use the survey findings as the basis for a published research article. But will the Medical Officer in charge agree to the survey? The lecturer suggests the Ministry of Health writes to the MO to explain the new Service Charter initiative and to invite his clinic to act as a pilot which will develop service standards based upon the voice of the local community. The letter must say clearly that the purpose of the pilot is for learning.

It takes time for the letter to be written but the MO in charge is surprisingly receptive when Sarah meets him again. He thinks if he is cooperative it could help his promotion prospects! Sarah organizes a community meeting which the young lecturer attends to decide the next steps. There is a voluntary organization which has been working in the community to provide support to families affected by HIV/AIDS. The community decides that this is the most trusted group to whom people would be willing to convey their views and experiences. It is agreed that the lecturer, Sarah's management committee and some members of the CSO will sit down to finalise the design of the questionnaire based upon some of those that the lecturer has seen in India. In this meeting, the group discusses how respondents will be selected and they agree that they must include the poor people in the community and women who have many young children. Sarah's committee and the CSO know where these people live. The group agrees that people should be interviewed in the homes away from the gaze of the clinic staff and that no-one will be forced to give an interview. The CSO will protect the confidentiality of respondents.

After the interviews have been conducted, the lecturer sits down with the interviewers to assist with the analysis of the results. Sarah and the lecturer agree that the results will be presented to the MO in charge and his staff, but not the Ministry of Health. After all, the MO is the person who must act on the data. The lecturer will prepare a separate report for the Ministry on the process which has been followed, highlighting how the report card will help in the development of a Service Charter which is a key reform intervention in the Ministry. A meeting is held between all parties at the clinic premises. The staff of the clinic are apprehensive initially, but Sarah explains that purpose of the survey is not to seek out and punish the guilty rather to work together to see how improvements can be made. The meeting goes well and the Medical Officer sees it as an opportunity to make a case to the central Ministry for additional staff and resources to improve services to the level demanded by the community.

To conclude the meeting, it is agreed that the findings of the report card will be used to develop the clinic's Service Charter. Sarah and the lecturer work together with the MO to develop the Charter. They use the information from the report card to identify a set of draft service standards for the clinic which will be included in the Charter. Once the draft is prepared they convene a workshop involving a cross-section of users and clinic staff to agree upon those standards which meet the expectations of client, which can be realistically achieved by the staff and resources of the clinic, and which are consistent with the minimum national standards. It is agreed that the Report Card will be repeated on an annual basis to monitor whether these standards are being met, and to highlight any staffing or capacity problems which cannot be addressed locally. These standards are published in the Service Charter, together with additional information on opening hours, fee schedules, a complaints mechanism and the responsibilities of clients. The Charter is visibly displayed in the registration area of the clinic.

Case study 4: Monitoring a National Public Administration Reform Strategy

The Government of (country x) has recently concluded an exercise to prepare a civil service management strategy. It has been produced by a local team of civil servants working under the Permanent Secretary in the Ministry of Public Service. The team is proud of their accomplishment and somewhat relieved that it is now over. It has taken nine months to carry out the diagnostic work and it has involved extensive consultation with a wide range of internal and external stakeholders. But what they do not realise is that the hard part – implementation – is about to start!

The strategy document is formally submitted to the Head of the Civil Service who establishes a small sub-committee of Permanent Secretaries to review it. This step takes everyone by surprise. The team and the Permanent Secretary, Public Service had expected that the Head of Civil Service would simply rubber stamp the document and then pass it on to the Cabinet for endorsement. But the Head of the Civil Service has previously been advised by the Permanent Secretary, Public Service, that she would be expected to chair a reform steering committee which should meet quarterly, at least in the initial stages of strategy implementation. This is what the donor agencies are pushing for. The Head of Civil Service asks herself: what will we do at these meetings? She thinks that one meeting should be devoted to deciding the reform priorities that will be supported in the next financial year. This meeting would need to be coordinated with the completion of the Government's annual budget process when allocation decisions are made. But what would happen in the other three meetings? Well, the committee could discuss some big issues of civil service management, but she does that with her PSs anyway at the regular monthly meetings. So she concludes that the other three meetings should concern themselves with monitoring the progress and impact of the PAR strategy.

Unfortunately, when she reviews the strategy document she cannot see what it is that her steering committee will be monitoring at their quarterly meetings. There is a brief technical section on "Monitoring and Evaluation" which talks about quarterly progress monitoring, a mid-term review and an impact evaluation. But she cannot find any indicators that would measure either the progress or impact of the reform. There is a beautifully crafted vision statement at the beginning of the document, but because it is very general, it is difficult to see how one would know whether it has been reached. Objectives have been crafted for the different components (e.g. restructuring of government, human resource management reform, pay reform, service delivery improvement) but these are very broad and not easy to measure. At the end of the document, there is a detailed action plan which sets out the tasks and activities that must be accomplished for each component on a day to day basis. She wonders, however, why a committee of Permanent Secretaries should be interested in whether job descriptions have been prepared for civil servants or whether ministries have surveyed the levels of satisfaction of their clients. Surely, the managers who are responsible for the different components should be monitoring these activities. All in all, she faces a dilemma. On the one hand, the high level objectives which she believes are important do not appear to be susceptible to measurement. On the other hand, the lower level tasks and activities which are easily measureable do not appear to be important.

She decides to ask UNDP's local governance adviser for help. Fortunately, the adviser has a good understanding of PAR measurement issues and the available tools. He knows that an "off the shelf" tool is inappropriate for monitoring a particular strategy and advises the Head of the Civil Service that government should develop a customised set of indicators. He adds that this work is timely because, once strategy implementation has started, they will have missed the opportunity to develop a set of baseline measures from which to measure improvements. UNDP offers to pay for an international consultant to work with the local team that developed the strategy in developing a monitoring framework.

When the international consultant arrives the team go back to the drawing board. They start by examining the high sounding vision statement for the future civil service and break it down into a number of constituent objectives related to its size, capability, accountability and incentives for civil servants. These four objectives are positioned alongside each other at the top of a white board. The international consultant then challenges the local team to clarify the reform strategy, after all they wrote the strategy document. He explains that they will adopt a "driver" model which identifies the key strategic priorities linked together in a causal fashion. This involves identifying the intermediate results that must be accomplished to achieve each objective and then specifying the reform actions to be carried out to achieve these results. The local team look puzzled initially because they had thought that it would just be necessary to define the indicators for each of the components. The consultant explains that the strategy map represents the pathway to achieving the four reform objectives. It defines the priorities, the relationship between them, and their sequencing. The local team realise that they have missed out this step and have jumped straight into the definition of the programme components. It is a common oversight.

Once they start to work on this exercise, the local team soon realise that they have all the information needed to develop the "driver" model in their heads. The international consultant is then able to facilitate this thinking process without imposing his views on the content of the strategy. They are conscious that the purpose of the exercise is to produce a valid set of indicators for the Steering Committee and they avoid defining the detailed steps for each reform intervention. Nonetheless, it takes two days to complete this exercise. The team then sits back and reviews its chart of "bubbles" and arrows. At this stage, the international consultant says they need to consider which information in the chart is likely to be required by the Steering Committee. They decide that the committee will wish to be informed at their quarterly meetings when key reform interventions have been completed, for instance when a revised civil service code of conduct has been developed. The indicators for the reform interventions are expressed simply as completion dates, namely milestones.

The final step is for the team to define the "indicators" for the intermediate results in the driver model and to collect baseline data against which progress can be monitored. The team has to take a pragmatic approach because they soon realise that in many instances it is simply too difficult to collect data for the ideal indicator. They end up using a number of proxy measures instead. Deciding on the indicators and collecting baseline measures turns out to be an iterative and time consuming process. After this experience, the team decides that it must define the procedures for collecting the data, the frequency with which it is required and the responsibilities for generating it. Otherwise, future measurement will simply not happen.

The baseline figures for each indicator are provided to the Head of the Civil Service and the team explain when improvement against each indicator can be expected. Clearly, no results can be achieved until after the relevant reform intervention has taken place. Based on the milestones established for the different reform interventions, the team advises the Head of the Civil Service that the Steering Committee should start to measure results at the end of the second year of implementation. The Head of Service is finally satisfied. The Steering Committee can keep track of the relevant milestones for the reform interventions on a quarterly basis and, once they have been reached, they can start to consider the impact of reform on an annual basis.

REFERENCES

Clay Wescott, 2008, World Bank Support for Public Financial Management: Conceptual Roots and Evidence of Impact, IEG World Bank

Nielson, 2007, Kiwis Count 2007, State Services Commission, December 2007

Lonti, Z and M. Woods, 2008, Towards Government at a Glance: Identification of Core Data and Issues related to Public Sector Efficiency, OECD Working Papers on Public Governance, No.7

Wouter Van Dooren, Nick Manning, Jana Malinska, Dirk-Jan Kraan, Miekatrien Sterk, Geert Boukaert, 2006, Issues in Output Measurement for "Government at a Glance", OECD Technical Paper 2

Oxford Policy Management, 2008, Assessing Public Financial Management Systems in Developing Countries: what works, what doesn't and why, OPM Briefing Notes

Jak Jabes, 2002, On the Futility of Governance Indicators: Lessons from Countries in Transition, Asian Development Bank

Daniel Kaufmann & Aart Kraay, 2007, On Measuring Governance: Framing Issues for Debate, World Bank

Steven van der Walle, 2005, Measuring Bureaucratic Quality in Governance Indicators

DAC Network on Governance, 2008, Donor Approaches to Governance Assessments: Sourcebook, OECD

Committee of Experts on Public Administration, 2005, Bottom-up approaches and methodologies to develop foundations and principles of public administration: the example of criteria-based organizational assessment, UN Economic and Social Council

Christopher Pollitt & Geert Bouckaert, 2000, Public Management Reform: A comparative Analysis, Oxford University Press

UNDP, Practice Note on Public Administration Reform

SOURCE GUIDE

Key:

ASSESSMENTS

1. Quantitative Service Delivery Surveys (QSDSs), by World Bank

1) **History:**
Many countries have large and ambitious public spending programs intended to address issues of efficiency and equity. Empirical evidence is severely lacking on questions of incentives and "moral hazard" despite a plethora of appealing theoretical arguments. The QSDSs help with assessing the implementation capacity of governments and the analysis of public expenditures and provide primary data on service providers for empirical research. QSDSs are a complementary tool to the PETS (see separate entry).

The first survey was carried out on the health sector in Uganda in 2000. The pilot study focused on Africa, but the techniques have been used in other country studies in health and education. Field testing of each survey is key to ensuring high quality results.

2) **Stated objectives:**
To examine the efficiency of public spending and incentives and various dimensions of service delivery in provider organizations, especially on the frontline.

3) **Measurement focus:**
QSDSs assess **service delivery**. They can be used for the analysis of public expenditure management reforms, service-delivery reforms, reforms to improve the efficiency of public expenditure, as well as cross-cutting public sector reforms.

4) **Information generated:**
While each QSDS is designed for the country situation, six core elements for all facility questionnaires have been identified:
- *Characteristics of the facility:* the size, ownership, years of operation, hours of operation, catchment population, competition from other service providers, access to infrastructure, utilities and other services, and range of services provided.
- *Inputs:* monetary values or quantities.
- *Outputs:* such as numbers of inpatient and outpatients treated, enrollment rates, and numbers of pupils completing final exams.
- *Quality:* such as staff behavior and composition, availability of crucial inputs, and provision of certain services, such as laboratory testing.
- *Financing:* sources of finance, amounts, and type (in-kind versus financial support).
- *Institutional mechanisms and accountability:* information on supervision visits, management structures, reporting and record-keeping practices, parent or patient involvement, and audits.

5) **Methodology:**
QSDSs are surveys of service providers. Data are collected both through interviews with managers and staff and from the service provider's records. In some cases, beneficiaries are also surveyed. Data are then triangulated to allow cross-validation of information. Research questions relate to incentives, provider behavior, cost-efficiency, and quality of service. Key stakeholders, including government agencies, donors and civil society organizations, are involved in the design.

6) Strengths & Weaknesses:
Strengths:
- Unique survey of actual service delivery.
- Use of hard data over judgments limits subjectivity.
- Users design the survey to fit their needs.
- Can be designed to be pro-poor and gender-sensitive.

Weaknesses:

• Data collection requires more effort, cost and time than perception surveys.
• Agents may have a strong incentive to misreport (or not report) key data, as the results of the data are used by donors in ways that benefit these agents.
• Necessary data to conduct a full QSDSs may not be available in all sectors in all countries at the facility level. Results suffer when data is limited, e.g. when service provision is not well recorded or is in-kind.

7) **Where to find it:**
For an overview, see http://go.worldbank.org/V68H7VFKG0 or http://go.worldbank.org/1KIMS4I3K0. For a detailed description, see http://povlibrary.worldbank.org/files/12933_chapter9.pdf

2. **Citizen Report Cards, by Public Affairs Centre, Bangalore, India**

1) **History: history, timeline, coverage, testing & validation**
Inspired by a private sector practice of conducting client satisfaction surveys, a small group of people in Bangalore, concerned about the city's deteriorating standards of public services, initiated an exercise in 1993 to collect feedback from users. The findings presented a quantitative measure of satisfaction and perceived levels of corruption, which mobilized citizen and government support for reform. Since then it has been replicated in many Indian cities as well as other countries. In 2006-07 the Public Affairs Centre developed a web-based self-learning course for those interested in implementing a citizen report card. The course was piloted in Pakistan and the Philippines and then revised before launching.

2) **Stated objectives:**
To provide public agencies with systematic feedback from users of public services.

3) **Measurement focus:**
Citizen Report Cards assess **service delivery**. They provide a rigorous basis and a proactive agenda for communities, civil society organizations or local governments to engage in a dialogue with service providers to improve the delivery of public services.

4) **Information generated:**
Questions relate to critical themes in the delivery of public services such as access to services, quality and reliability of services, problems encountered by users of services and responsiveness of service providers in addressing these problems, transparency in service provisions like disclosure of service quality standards and norms, and costs incurred in using a service including hidden costs such as bribes.

Example of a question:
• "Did you have to make any payments related to the delivery?"
   1- Yes (go to next question)
   2- No (skip next question)
• "How much did you pay?" _____

5) **Methodology:**
Citizen Report Cards are surveys of feedback from users of public services on these services. Data are collected through interviews with users. Results are published and a summative satisfaction score is calculated that captures the totality of critical service-related parameters. Preliminary focus group discussions with citizens from the population of interest are used to help finalize the purpose of the report card and to design the survey instrument.

6) **Strengths & Weaknesses:**
Strengths:
• Contact with actual users ensures first-hand information.
• Users design the survey to fit their needs.
• Maximizes citizen ownership due to the key role they play in implementation.
• Can be designed to be pro-poor and gender-sensitive.

Weaknesses:
- Requires expertise on quantitative methods and statistics that may not exist in-house.
- When the means for citizens to voice their opinion (through elections, public meetings, and other ways) are less, or absent, political institutions and processes are rarely designed to use feedback from citizens.
- Where individuals fear punishment for commenting on the government, the reliability of collected feedback is in question.

7) **Where to find it:**
The Public Affairs Centre research page, with links to past report cards, is at http://www.pacindia.org/issues/research. Their Citizen Report Card Learning Toolkit is at http://www.citizenreportcard.com.

3. **Common Assessment Framework (CAF), by European Institute of Public Administration**

1) **History:**
The CAF is a result of cooperation among the EU ministers responsible for Public Administration. A pilot version was presented in May 2000 and a first revised version was launched in 2002. Between 2000 and 2005 about 900 European public administrations used the CAF to improve their organisations. It has also been applied outside Europe. The CAF was revised in 2006.

2) **Stated objectives:**
To assist public-sector organisations across Europe in using quality management techniques to improve performance.

3) **Measurement focus:**
The CAF assesses **leadership, civil service management, public financial management, policy-making system** and **service delivery.** It is designed for use in all parts of the public sector at the national/federal, regional and local level. It may be used under a wide variety of circumstances, e.g. as part of a systematic programme of reform or as a basis for targeting improvement efforts in specific public-service organisations. In may also be used in part of an organisation, e.g. in a selected section or department.

4) **Information generated:**
Enablers criteria include the following:
- Leadership addresses the way in which leaders develop, implement, and monitor the organisation's management system and review performance and results.
- Strategy and planning addresses the way the organisation implements its mission and vision and translates this into plans, objectives and measurable targets.
- People addresses the way in which employees interact with each other and manage the available resources.
- Partnerships and resources addresses how the organisation plans and manages its key partnerships – especially with citizens/customers – in order to support its strategy and planning and the effective operation of its processes.
- Processes addresses how the organisation identifies, manages, improves and develops its key processes in order to support strategy and planning.

Results criteria include the following:
- *Citizen/customer-oriented results* addresses the results the organisation is achieving in relation to the satisfaction of its citizens/customers with the organisation and the products/services it provides.
- *People results* addresses the results the organisation is achieving in relation to the competence, motivation, satisfaction and performance of its people.
- *Society results* addresses the results the organisation is achieving in satisfying the needs and the expectations of the local, national and international community.
- *Key performance results* addresses the results the organisation is achieving with regard to its strategy and planning related to the needs and demands of the different stakeholders (i.e. external results); and the results the organisation has achieved in relation to its management and improvement (internal results).

Each criterion is divided into sub-criteria. For example, the sub-criteria for People are:
- Plan, manage and improve human resources transparently with regard to strategy and planning
- Identify, develop and use competencies of employees, aligning individual and organisational goals
- Involve employees by developing open dialogue and empowerment

Under the sub-criteria are suggestions for indicators. Some of these are gender-sensitive. For example:
- The Citizen/customer-oriented results criterion suggests indicators of complying with diversity and gender aspects.
- The Partnerships and resources criterion suggests indicators for introduction of innovative systems of budgetary and cost planning, including gender budgets.

5) **Methodology:**

The CAF is a representation of all aspects that must be present in the proper management of an organisation in order to achieve satisfactory results. All these elements are translated into nine criteria, divided into subcriteria. The first five criteria deal with the enabler features of an organisation: what the organisation does and how it approaches its tasks to achieve the desired results. The last four criteria deal with the results achieved in the fields of citizens/customers, people, society and key performance.

A group from within the organisation evaluates the organisation. The self-assessment group should be as representative of the organisation as possible. Each member of the self-assessment group gives an accurate assessment of each subcriterion based on their own knowledge and experience of working in the organisation. Then the group meets to agree on the strengths, areas for improvement and the scores on each subcriterion. Scoring is done on the basis of one of two methods, depending on the time available and the organisation's experience and maturity.

Broad consultation involving management, internal stakeholders, and external stakeholders when applicable is recommended in the development of the assessment tool.

6) **Strengths & Weaknesses:**

Strengths:
- Well-suited for helping to generate reform plans that are tailored to the organisation and have internal buy-in.
- A "light" model, easy to use and especially suited to gaining an initial impression in a short period of time of how an organisation performs.
- Some customization is recommended, while compulsory basic elements allow for comparison.
- Available in a range of European languages.
- Some gender-sensitive examples provided for ways to judge sub-criteria.

Weaknesses:
- Free registration is required in order to use the method online.
- Requires an experienced facilitator when used in an organization for the first time.
- Action plans are likely to focus on training, which is attractive and non-contentious, rather than tackling the most difficult problems.
- Minimal pro-poor information, although each organization can customize as desired.

7) **Where to find it:**

Homepage at http://www.eipa.eu/en/topics/show/&tid=191. English-language brochure with method at http://www.eipa.eu/files/File/CAF/Brochure2006/English_2006.pdf

4. **Country Governance Assessment (CGA), by Public Administration and Public Financial Management sections Asian Development Bank (ADB)**

1) **History:**

The quality of governance in ADB's developing member countries represents both a major risk and an opportunity for ADB in pursuing its central objective of poverty reduction. ADB has always collected and made use of information about governance in these countries . However, since a formal governance policy was adopted in 1995, governance work has significantly increased its profile, generating a need for more systematic study of governance conditions and issues.

CGAs have been completed in various forms since 1997. In 2002 the method was reviewed and evaluated for its use and impact; recommendations made for how to carry out further governance analysis to complement but not replace the CGAs. Since 2003, 10 reports have been published for Asian countries, and others are used internally.

2) **Stated objectives:**
To systematically assess the quality of governance of member countries and to strengthen the linkage between the quality of governance and levels and composition of assistance.

3) **Measurement focus:**
The Public Administration section of the CGA assesses **policy-making system** and **civil service management**. The Public Financial Management section assesses **public financial management**.

CGAs are used primarily by the ADB, with secondary use by the assessed governments. Governance information and analysis is used for identifying the impact of governance on the poor; conducting policy dialogue with assessed governments on key governance issues; guiding public discussion, information-sharing, or joint projects on country governance with other stakeholders (e.g., donor agencies or civil society organizations); developing strategies and programs to address key governance issues; reaching decisions on the size and composition of ADB's country portfolios; identifying risks and opportunities from the governance environment in sectors and projects; developing indicators and benchmarks of governance at various levels; identifying the relationship between public governance and the private sector or civil society.

4) **Information generated:**
Under the Public administration section, the following sub-sections are covered:
- Government system: To assess the overall policy making and implementation framework.
- Civil service: To assess the degree to which civil service arrangements are implemented in line with existing laws and to determine whether human resources are managed with reference to accepted public management practices.
- Local Governance: To evaluate the capacity of local governments to respond adequately to the real needs of their constituencies.

Under the Public financial management section, the following sub-sections are covered:
- Revenue administration
- Public expenditure management
- Public sector accounting and internal control
- Public sector auditing
- Inter-governmental fiscal relations
- Specific anti-corruption efforts

Each sub-section contains a series of guide questions. For example, under the Civil service sub-section are questions on the following topics:
- Legality and accountability of the civil servants
- Impartiality, and integrity of civil servants
- Management of the public service and human resources management issues

5) **Methodology:**
CGAs are narrative reports based on the CGA Framework, a general method divided into sections that contain guidance questions for evaluation. There is no set of rules on how a governance analysis might be designed, managed, or phased. The CGA Framework provides guidance on the analytical basis for a CGA but not on process, evidence acquisition, or resources. The result has been a good deal of variation in both CGA process and product. Broadly speaking, task managers from ADB have adopted three approaches:
(i) Brief reports by staff based largely on analyzing existing documents on country governance, supplemented by short visits to the country;
(ii) Wide-ranging desk research with field visits, a basic level of participation (in the form of interviews and consultation), and selected use of national or international consultants; and
(iii) Participation by government and country stakeholders in the design and execution of the analysis, often involving extensive consultation and workshops.

6) **Strengths & Weaknesses:**

Strengths:
- Sound assessment method that has been tested and reviewed.
- Focus on qualitative data emphasizes full understanding of topics and not simply "a number".
- Loose method allows for customization, while standard framework allows for some comparison.

Weaknesses:
- A limited budget will mean a decision between greater scope or depth of coverage.
- Yet to be implemented outside the context of the ADB.
- Not designed to be used by public servants, and therefore the information may not be useful for their purposes.
- Not pro-poor. ADB recommends separate poverty analysis of all countries.
- Not explicitly gender-sensitive, although the method could be further developed to include gender-sensitive questions.

7) **Where to find it:**

http://www.adb.org/Governance/gov_cga.asp

5. Capability Reviews, by UK Civil Service

1) **History:**

In October 2005, the UK Civil Service began development of Capability Reviews to both assess how well equipped departments are to meet their delivery challenges and provide targeted support to make any improvements required. The programme targets underlying capability issues that have an impact on effective delivery. The Capability Review programme is part of the wider reform of the civil service. The reports were published in July 2006. More than 15 have been published since then.

2) **Stated objectives:**

To assess how well equipped UK government departments are to meet their delivery challenges and provide targeted support to make any improvements required.

3) **Measurement focus:**

Capability Reviews assess **leadership**. They are used to inform reform by the Civil Service and the departments under review. The public may also use the reviews to better understand public institutions and to advocate for change.

4) **Information generated:**

The model of capability that forms the basis for the capability reviews is divided into three areas: Leadership, Strategy and Delivery. These in turn comprise sub-areas as follows:

Leadership
- Set direction: how planning is done
- Ignite passion, pace and drive: whether and how leaders set an example
- Take responsibility for leading delivery and change: whether and how change is pushed through
- Build capability: how employee performance is managed

Strategy
- Focus on outcomes: features of the strategy
- Base choices on evidence: how strategic choices are made
- Build common purpose: how work with partners is achieved

Delivery
- Plan, resource and prioritise: features of delivery plans
- Develop clear roles, responsibilities and business model(s): how division of labor is determined
- Manage performance: how performance is evaluated and built upon

5) Methodology:
Capability reviews are narrative reports paired with ratings. The reviews are conducted using a standard model of capability developed by the Capability Reviews Team through consultation with senior civil service leaders and external experts. The reviews are conducted by teams including business and public sector leaders from outside the government. The conclusions are evidence based and the findings are reviewed for consistency by an external moderating panel.

6) Strengths & Weaknesses:
Strengths:
- External review using a standard method but intended for internal use.
- Easily understood method that can be applied by experienced managers and civil servants.
- Central oversight of implementation by Cabinet Office Capability Review Team helps ensure quality and comparability.
- Applicable to any department throughout the civil service.

Weaknesses:
- Would likely require modification for country contexts outside the UK.

7) Where to find it:
http://www.civilservice.gov.uk/about/accountability/capability/index.asp

6. Public Expenditure Tracking Surveys (PETSs), by World Bank

1) History:
Government resources often pass through several layers of government bureaucracy (and the banking system) before reaching service facilities, which are charged with the responsibility of exercising the spending. Information on actual public spending is seldom available in developing countries. A PETS – frequently carried out as part of a public expenditure review – tracks the flow of resources through these strata in order to determine how much of the originally allocated resources reach each level. PETSs are a complementary tool to the QSDSs (see separate entry).

The first survey was carried out on the education sector in Uganda in 1996; PETSs have since been carried out worldwide in health and education. Field testing of each survey is key to ensuring high quality results.

2) Stated objectives:
To track the effect of public expenditure on growth and/or social outcomes.

3) Measurement focus:
PETSs assess **public financial management**. They can be used for the analysis of public expenditure management reforms, reforms to improve the efficiency of public expenditure, cross-cutting public sector reforms, anti-corruption, and service delivery reforms.

4) Information generated:
While each PETS is designed for the country situation, six core elements for all facility questionnaires have been identified:
- *Characteristics of the facility*: the size, ownership, years of operation, hours of operation, catchment population, competition from other service providers, access to infrastructure, utilities and other services, and range of services provided.
- *Inputs:* monetary values or quantities.
- Outputs: such as numbers of inpatient and outpatients treated, enrollment rates, and numbers of pupils completing final exams.
- *Quality:* such as staff behavior and composition, availability of crucial inputs, and provision of certain services, such as laboratory testing.
- *Financing:* sources of finance, amounts, and type (in-kind versus financial support).
- *Institutional mechanisms and accountability:* information on supervision visits, management structures, reporting and record-keeping practices, parent or patient involvement, and audits.

Some sample questions appear in Table 1.

Table 1: Sample questions for a PETS on the education sector (designated teacher has been selected through sampling)

Question	Unit
How much does this teacher receive in salary each month?	currency figure
How much is deducted from each payslip automatically?	currency figure
Who pays the teacher's salary?	1 = Nat'l gov't; 2 = school; 3 = community; 4 = other
Prices and wages vary considerably across different parts of this country. What is a typical hourly wage for a manual laborer in this area?	Currency figure
Do you think it is possible to support a family only on the salary that this teacher earns?	1 = Yes; 2 = no

Source: Ritva Reinikka and Nathanael Smith, 'Public expenditure tracking surveys in education,' International Institute for Educational Planning, 2004.

5) **Methodology:**

PETSs are surveys of service providers. Data are collected both through interviews with managers and staff and from the service provider's records. In some cases, beneficiaries are also surveyed. Data are then triangulated to allow cross-validation of information. A PETS focuses on service provider behavior, incentives, and relationship between providers, policy-makers and users. Key stakeholders, including government agencies, donors and civil society organizations, are involved in the design.

6) **Strengths & Weaknesses:**

Strengths:
- Unique analysis of the flow of resources through many layers of bureaucracy that can help identify the location and extent of obstacles to resource flows (financial, staff, equipment).
- Quantitative data on service delivery that limits subjectivity.
- Users design the survey to fit their needs.
- Can be designed to be pro-poor and gender-sensitive.
- Limited cost and time required in relation to its objectives.

Weaknesses:
- Agents may have a strong incentive to misreport (or not report) key data.
- Results suffer when data is limited, e.g. when service provision is not well recorded or is in-kind.

7) **Where to find it:**

For an overview, see http://go.worldbank.org/AGLWH0RV40 or http://go.worldbank.org/1KIMS4I3K0. For a detailed description, see http://povlibrary.worldbank.org/files/12933_chapter9.pdf

7. Self-Assessment Tool for Customer Service Excellence, by UK Civil Service

1) **History:**

Customer Service Excellence is the UK's new formal assessment standard for customer service, launched by the civil service in 2008. It is a result of a detailed review conducted in 2005-06 that recommended that a new customer service standard, based on the key drivers of satisfaction, should be developed to replace Charter Mark, the previous standard. It was developed to offer public services a practical tool for driving customer-focused change within their organisation. The foundation of the tool is the UK government's Customer Service Excellence standard.

2) **Stated objectives:**
   To assess an organization's capability, identify areas for improvement, and formulate a plan to move forward.

3) **Measurement focus:**
   Customer Service Excellence assess service delivery. It is designed for UK organizations in the public, private or voluntary sectors to identify areas and methods for improvement, to explore and acquire new skills in the area of customer focus and customer engagement, and to demonstrate their competence.

4) **Information generated:**
   Customer Service Excellence is divided into five criteria, which contain subcriteria divided into different elements with guidance. The five criteria of Customer Service Excellence are:
   - Customer Insight: Effectively identifying your customers, consulting them in a meaningful way and efficiently measuring the outcomes of your service.
   - The Culture of the Organisation: A truly customer-focused culture built through a commitment to it throughout an organisation.
   - Information and Access: Accurate and comprehensive information is delivered or available through the most appropriate channel for customers.
   - Delivery: How you carry out your business, the outcomes for your customer, and how you manage any problems that arise.
   - Timeliness and Quality of Service: The promptness of initial contact and keeping to agreed timescales, combined with quality of service to ensure the best possible result for customers.

   The sub-criteria for Achieved Timely Delivery under Timeliness and Quality of Service contains the following elements and guidance:

Element	Guidance
We monitor our performance against standards for timeliness and quality of customer service and we take action if problems are identified.	• Methods used for monitoring performance in relation to standards for timeliness and quality of customer service. • Steps taken to address timeliness of delivery where this does not meet customer expectations.
We are meeting our current standards for timeliness and quality of customer service and we publicise our performance against these standards.	• Current actual performance against all standards and targets for timeliness and quality of service and evidence of publication. • Customer feedback on achievement of service delivery within agreed timescales. • Examples of 'mystery shopper' evidence, where relevant.
Our performance in relation to timeliness and quality of service compares well with that of similar organisations.	• Comparative performance data relating to timeliness and quality of customer service obtained from benchmarking exercises.

One element addresses disadvantaged groups, under Customer Insight:

Element	Guidance
We make particular efforts to identify hard-to-reach and disadvantaged groups and individuals and have developed our services in response to their specific needs.	• Methods used to consult and involve hard-to-reach and disadvantaged groups and individuals, particularly people who have a number of different needs and customers whose views you rarely hear. • Examples of how you identified needs that you were not meeting and the changes you made to avoid excluding or under-representing certain people.

5) **Methodology:**
The Customer Service Excellence standard is an assessment questionnaire. Based on the elements it contains, evidence is gathered and scores are assigned on a 1 to 4 scale. Individuals or teams from within the assessed organization carry out the assessment based on available data in the organization. Evidence can be formal or informal.

6) **Strengths & Weaknesses:**
   • Thorough self-assessment tool based on recognized standards.
   • Well-suited for helping to generate reform plans that are tailored to the organization and have internal buy-in.
   • Limited resources required.

   Weaknesses:
   • Self-assessment, while important, leaves open the possibility of inflating results.
   • Registration is required in order to use the method online, and only UK organizations can register.
   • Minimal gender-sensitive or pro-poor information.

7 **Where to find it:**
Tool website at http://www.cse.cabinetoffice.gov.uk/welcomePage.do. Method available at http://www.cse.cabinetoffice.gov.uk/UserFiles/Customer_Service_Excellence_standard.pdf

8. **Performance Measurement Framework, by Public Expenditure and Financial Accountability (PEFA)**

1) **History:**
The PEFA program launched the Performance Measurement Framework in June 2005 as part of 'the Strengthened Approach to Supporting PFM Reform'. The original framework was tested in draft form in at least 24 countries and revised before the final version was released. Consultations were held with technical experts from PEFA partners, professional organizations, other donors and government practitioners from 17 African countries. In 2006, a report on monitoring the application of the Performance Measurement Framework was prepared. A follow-up monitoring exercise was undertaken during 2007. The framework has been used to evaluate a wide range of aid-recipient countries worldwide. Assessments can be of the national and sub-national level.

2) **Stated objectives:**
To assess and develop essential public financial management (PFM) systems by providing a common pool of information for measurement and monitoring of PFM performance progress, and a common platform for dialogue.

3) **Measurement focus:**
The Performance Measurement Framework assesses **public financial management.** The intended uses of the applications have usually fallen in three categories: (i) to monitor the impact of PFM reforms, (ii) to feed into the dialogue between government and donors on the scope, priorities and sequencing of the government's PFM reform programs or adjustments thereof, (iii) to inform donors on fiduciary issues related to their lending or grant aid programs, particularly in relation to budget support operations.

4) **Information generated:**
The Performance Measurement Framework comprises a set of indicators in four areas:
* Credibility of the budget
* Comprehensiveness and Transparency
* Budget Cycle
* Donor Practices

Under these are sets of performance indicators. For example, under Credibility of the budget are:

* Aggregate expenditure out-turn compared to original approved budget
* Composition of expenditure out-turn compared to original approved budget
* Aggregate revenue out-turn compared to original approved budget
* Stock and monitoring of expenditure payment arrears

5) **Methodology:**
The Performance Measurement Framework comprises a PFM performance report and a set of high level indicators that draw on international standards. These are used to produce a narrative report and scores. The framework may be implemented by the government alone (self-assessment), by the donors alone, by government and donors collaboratively, or by an independent group. In practice, most of the assessments have included participation by the partner government, the involved donor agencies and international consultants. Barely ever is civil society included. The assessment team has typically comprised economists, accountants, audit experts and procurement experts. Previous analytical work has typically been used as an initial source of information for the assessors, followed by interviews with officials from the assessed government and its agencies. Views of nongovernment actors (private sector and civil society organizations) have rarely been consistently included. Some past assessments have included external review.

6) **Strengths & Weaknesses:**
Strengths:
* Well-researched method for PFM assessment based on international standards, with broad donor buy-in.
* Combination of narrative and scores provides explanation behind the numbers.
* Limited guidance allows for flexible implementation.

Weaknesses:
- Implementation costs can be high (USD 100,000 or more)
- Framework calls for little CSO participation.
- Despite goal of poverty-reduction, indicators are not pro-poor.

7) **Where to find it:**
http://www.pefa.org/index.php

9. Public Officials' Survey, by World Bank-Netherlands Partnership Program (BNPP)

1) **History:**
In 1999-2001, an analytical framework was used to design a series of surveys of the views of public officials concerning their institutional environment and to analyze the information generated. The need for an analytical framework emerged from the lessons of past experience in the World Bank, which documented that during 1980-97, only one-third of the Bank's closed civil service management interventions had successful outcomes. The framework rests on the premise that a public official's performance is a function of the incentives provided by their institutional environment. Surveys were designed to elicit a contextual description of the public sector. Fifteen surveys have been completed worldwide. Each survey was pre-tested before implementation.

2) **Stated objectives:**
To map the strengths and weaknesses of the public sector and identify potential pay-offs with reform interventions.

3) **Measurement focus:**
BNPP surveys assess **civil service management** and **public financial management.** Their use has not been evaluated, but they provide information relevant to those interested in public administration reform.

4) **Information generated:**
There is no common questionnaire for the surveys. However, the conceptual framework for the surveys divides indicators into two sections each with three dimensions:

Indicators for measuring institutional environment
- Rule credibility
- Policy credibility
- Resource adequacy and predictability

Indicators for measuring performance
- Results-focus
- Accountability
- Employee morale

Some examples of past indicators are as follows:

Within an organization, different people are hired for different reasons. This survey is interested in learning about general hiring trends. In your organization would you say that employees were hired more often for:
1. *good performance at entry*
2. *test qualifications*
3. *met job requirements*
4. *personal connections*
5. *political connections*
6. *family status in community*
7. *payment of gift to some*
8. *public authorities*
9. *knew a senior official*

For each of the following aspects of budget management, does your organization have formal, written guidelines describing how these functions are to be carried out. If so, are they carried out in full, in part, or not at all? Please answer on a scale of 1 to 4, where 1 corresponds to "not at all" and 4 corresponds to "yes, in full"

- *budget preparation*
- *procurement*
- *budget execution*
- *audit*
- *production of final end-of-year accounts*

Do you expect that actual funds will differ from budgeted funds by a large amount or a small amount?
1=a small amount; 4=a large amount

Some people consider that emigration of educated and trained persons from the Caribbean to US, UK and Canada drains the public sector of skilled persons. How severe is this problem in your organization?

1. *Very severe*
2. *Severe*
3. *Moderate*
4. *Mild*

How many employees of your organization with excellent performance have not received recognition/reward?

- *Many*
- *Some*
- *A few*
- *No one has received recognition*

5) **Methodology:**

BNPP surveys are surveys of public officials. The surveys seek a country-specific approach to distinguishing between groups of officials. In the case of Bolivia, for example, officials from central administrations, decentralized institutions, and regional administrations were surveyed and the analysis distinguished between consultants and other public officials. In Bangladesh, the analysis distinguished between officials of ministries, autonomous corporations and district-level officials. In the East Caribbean states, perceptions of managers, middle-level officials and general officials were differentiated. Focus groups were used to pre-test questionnaires, but otherwise surveys were designed by World Bank staff. Answers were converted to numbers to determine a final score for each set of indicators.

6) **Strengths & weaknesses:**

Strengths:
- Public officials surveys that go beyond best-practice models to examine the incentive structure in the public sector.
- Broad framework allows for customization while providing some guidance and structure.
- Can be designed to be pro-poor and gender-sensitive, although no guidance is provided.

Weaknesses:
- Lack of guidance not suitable for potential implementers unfamiliar with survey design; the method is complex and not user-friendly.
- Some risk of false reporting, especially on questions related to corruption and other poor behavior – should be triangulated with other data.
- No use or update since 2001.
- Non-participatory approach, although could be adapted.

7) **Where to find it:**

Homepage at http://go.worldbank.org/7E2U3B2VC0. Conceptual framework is available at http://www1.worldbank.org/publicsector/civilservice/wps2427.pdf.

10. **Country Assessment in Accountability and Transparency (CONTACT), by UN Development Programme (UNDP)**

1) **History:**

A need was felt among the international donor community to have a comprehensive tool to assess, within the overall governance picture, the financial architecture of a country including its integrity systems. With this in mind, UNDP published this guide in 2001. It was developed jointly with various UN agencies and the World Bank. It was designed as a dynamic and continuously evolving instrument that would require constant update through feedback based on best practices and pilot testing. However, no further documents have been published under this title.

2) **Stated objectives:**

To assist governments in conducting a self-assessment of their financial management and anti-corruption systems.

3) **Measurement focus:**

CONTACT assesses **public financial management**. It can be used by governments seeking to assess the strengths, weaknesses, needs and priorities of their country's financial accountability and integrity improvement systems and processes. It can also be used by consultants hired by donors to assess the financial accountability and integrity improvement systems and processes of a country.

4) **Information generated:**

CONTACT provides self-assessment checklists in financial accountability in the following areas:
- accounting infrastructure
- information management
- expenditures and budgeting
- internal control and internal auditing
- financial reporting
- external auditing
- revenue administration
- debt management
- project and foreign aid management
- procurement and assets management
- integrity improvements to prevent and control corruption
- public sector cash management

Some examples of checklist questions are:
- Is there a clear and comprehensive legal basis for accounting and audit in the public sector, such as the Constitution, Finance Act (or Financial Administration and Audit Act), Appropriation Acts, Treasury Regulations (or Instructions), Local Authorities Finance Act, Public Debt Act, enabling acts for public corporations?
- Does the [financial] plan show allocations to major programs (primary health care, secondary education, information technology, etc)?
- Taking the policies and regulations as a whole, can it be stated that an adequate framework has been established for an adequate Internal Control system to flourish?
- Are the duties and obligations of each revenue authority properly and explicitly laid down in the revenue legislation?

5) **Methodology:**

CONTACT is a set of generic guidelines with accompanying assessment checklists. It was designed to assist governments on an ad hoc basis and not as part of any program for implementation. No guidance is provided for how the checklists should be evaluated or what form the results should appear in.

6) **Strengths & weaknesses:**

Strengths:
- Well-researched, extensive list of indicators for financial management assessment.
- Designed as a guide that can be used according to the needs of the implementer.

Weaknesses:
- Not designed as a method, therefore would require considerable adaptation and focus in order to be used effectively.
- Many questions require subjective judgment without guidance as to how to ensure objectivity.
- Only minimal mention of pro-poor or gender-sensitive aspects.

7) **Where to find it:**
Homepage at http://www.undp.org/governance/docs/AC_guides_contact.htm. Complete guide at http://www.undp.org/governance/docs/AC_Guides_CONTACT2001.pdf.

11. Evaluation Matrix of Civil Service Human Resource Management in the European Union, by World Bank

1) **History:**
This report was part of a regional initiative undertaken by the World Bank to study the impact of European Union accession on World Bank borrowers in Central and Eastern Europe. The report reflects field analyses carried out through several missions to Estonia, Hungary, and the Czech Republic. Using international best practice as a reference, an evaluation matrix was developed to provide a specific assessment of progress on institutional reforms. The study assesses current administrative capacity and projects future requirements not just for EU accession but for broader qualification as a modern, competitive public administration on a global basis. Standards were discussed and tested with public administration experts from a range of OECD countries. The report was published in 2000.

2) **Stated objectives:**
To help countries specify targets in order to attain the level of administrative development necessary both for EU accession and global competitiveness.

3) **Measurement focus:**
The matrix assesses civil service management and the policy-making system. It was intended for EU candidate governments seeking to specify targets for their public sector.

4) **Information generated:**
The matrix is divided into three administrative issue areas that are further categorized into specific dimensions or objectives that are associated with well-performing, advanced country public administrations:

Civil Service Human Resource Management
- Legal and ethical framework
- Institutional framework
- Employment and pay policy and management
- Human resources management policies and practices
- Training and career development
- Management practices and culture

Policy Formulation and Coordination
- Institutional arrangements
- Policy processes and procedures
- Human resources capacity

EU Accession Management
- EU accession strategy
- Institutional arrangements
- Management processes and procedures
- Human resources issues

For each of these areas is provided good practice criteria, performance indicators used and benchmark information, and other optimal indicators.

5) **Methodology:**
The matrix is designed for self- or external assessment. Judgments were reached on individual country assessments for each criterion in the matrix. World Bank staff and country officials assessed performance individually and jointly as a team, reaching a consensus on each variable. Each criterion was then assigned a numerical score between 0 and 5. A qualitative summary as well as notes on action required were also provided.

6) **Strengths & weaknesses:**
Strengths:
- Well-researched and wide-ranging tool for civil service evaluation, despite its purpose related to EU accession.
- Broad guidelines mean that it can be used to suit the purposes of the implementer.
Weaknesses:
- Design and testing in European countries as well as lack of updating could mean revisions are required.
- Emphasis on laws and regulations, reflecting legalistic traditions in Continental Europe, rather than practice on the ground.
- Contents of the matrix are not pro-poor or gender-sensitive.

7) **Where to find it:**
The report is available at http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2000/08/14/000094946_00072405371429/Rendered/PDF/multi_page.pdf. The matrix is on pages 23 to 39.

12. **Control and Management System Baselines, by SIGMA (joint initiative of the OECD and the EU)**

1) **History:**
The control and management system baselines were prepared by the SIGMA Secretariat upon the request of the European Commission in close co-operation with various directorate generals of the Commission, as well as the European Court of Audit (in the area of external audit). SIGMA was to revise the baselines at regular intervals to keep pace with new EU legislation and developments. However, new baselines have not been published since 1999. The baselines and the work related to them support efforts by the European Commission in the area of EU enlargement.

2) **Stated objectives:**
To provide a basis for conducting assessments of central management and control systems to support efforts by the European Commission in the area of EU enlargement.

3) **Measurement focus:**
The baselines assess **civil service management, public financial management**, and **the policy-making system**. They are used by the European Commission for assessing EU candidate countries and other countries of interest.

4) **Information generated:**
There are 6 control and management system baselines:
- Public sector financial control baseline
- Public procurement management systems baseline
- Public expenditure management systems baseline
- Public sector external audit baseline
- Civil service baseline
- Policy-making and co-ordination machinery baseline

They vary in format and types of content, from a series of questions to more narrative form. Some examples are:

There should be an effective central administrative body with the capacity and authority:
- *to ensure that co-ordination arrangements are enforced;*
- *to provide adequate logistical support to the centre of decision-making;*
- *to ensure the recording and circulation of decisions;*
- *to monitor the implementation of decisions (including the necessary secondary legal instruments).*

There should be a clear legal basis for the establishment of a Public Procurement Office with overall responsibility for the design and implementation of public procurement policy. Such an office may be located in one Ministry or may be under the authority of the Prime Minister or the Council of Ministers or the Parliament.

Budget and accounting categories at the national level should have a common classification that facilitates policy analysis and promotes accountability. National accounting concepts need to be made compatible with concepts related to the disbursement of EC funds – commitments, payments, eligible expenditures, etc. Fiscal reporting should be timely, comprehensive, reliable, and identify deviations from the budget. Procedures for evaluating the efficiency and effectiveness of expenditure policies and programmes – including those funded from EC sources – should be established.

Are the SAIs annual and other reports prepared in a fair, factual and timely manner?

5) **Methodology:**
The baselines are assessment tools, either in the form of a questionnaire or a narrative description. They are used to generate narrative reports. Sigma staff carry out the assessments through interviews with government authorities, information from the EU and other donors, foreign and local public analyses, and work with other sectors of the OECD.

6) **Strengths & weaknesses:**
Strengths:
- A recognized standard used by practitioners.
- Wide-ranging coverage of public administration reform issues designed by specialists.
- Solid foundation from which to draw standards for public sector assessment.
- Modular design can be applied flexibly according to context.
Weaknesses:
- Designed specifically for the EU, based on good practice in EU member states and EU regulations and directives, and would therefore require some adaptation for other contexts.
- Some indicators are vague and/or subjective and would be difficult to evaluate without further guidance than is provided.
- Emphasis on implementation by experts precludes government buy-in through early engagement.
- Non-participatory method.

7) **Where to find it:**
See http://www.oecd.org/dataoecd/57/32/35007180.pdf.

13. Human Resources Self-Assessment Guide, by State of Texas

1) **History:**
Staff of the Texas State Auditor developed the latest version of the Human Resource Management Self-Assessment Guide, first developed in 1998. The methodology was developed by collecting and analyzing information on human resource criteria, state and federal statutes, and current business practices. It is designed for use in Texas state universities and agencies.

2) **Stated objectives:**
To provide formal guidance to state agencies and universities in assessing the effectiveness of human resource management controls, and to provide helpful resources to improve performance.

3) **Measurement focus:**
The Human Resources Self-Assessment Guide assesses **civil service management**. It is used by Texas state universities and agencies to determine human resources management strengths and weaknesses.

4) **Information generated:**
The Human Resources Self-Assessment Guide is divided into six Self-Assessment Modules, each with a separate questionnaire:

- Human Resources Management Practices
- Recruitment and Selection
- Training and Development
- Compensation and Benefits
- Employee Relations
- Safety, Health, and Security

Examples of questions include:

Does your organization have a policy or procedure for ensuring that employees who are transferred to another division/department or whose supervisors leave are regularly evaluated?
- *Yes*
- *Yes, in draft form*
- *No*

When does your organization report the number of minorities and women employed by your organization?
- *Not later than the seventh workday of each calendar year*
- *Within 60 days of the close of the fiscal year*
- *Within one year of the close of the fiscal year*
- *Upon request*
- *Never*

Does your organization's human resources department have controls in place to ensure that supervisors and managers consult with a human resources representative before employees are suspended, reprimanded, or fired?
- *Yes*
- *No*

5) **Methodology:**
The Human Resources Self-Assessment Guide is a questionnaire that can be accessed and completed online. After completion a score is automatically calculated. Questions are answered based on a review of documentation maintained by the assessed organization and in consideration of organization-wide human resources practices. Anyone can access the questionnaire and carry out the assessment based on any technique they choose. Results are not recorded in any central location.

6) **Strengths & weaknesses:**
Strengths:
- Broad, publicly accessible human resources questionnaire in convenient online format; despite some basis in Texas and U.S. law, most questions are widely applicable.
- Some gender-sensitive and pro-poor indicators.
- Limited resources required.
Weaknesses:
- Self-assessment, while important, leaves open the possibility of inflating results.
- Method for calculation of scores not publicly available.
- While the method is based in some recognized good practices, adaptation to the local context would be advisable before implementation in new areas.
- Emphasis on framework and regulations rather than practice.

Where to find it:
Homepage at http://www.hr.state.tx.us/apps/hrsag.

14. Human Resource Management (HRM) Assessment Instrument, by Management Sciences for Health

1) **History:**
The HRM Assessment Instrument is based on the Institutional Development Framework developed by the Family Planning Management Development project of Management Sciences for Health in 1998. It was pilot-tested with NGOs in Uganda, Zambia, Albania, and Bolivia.

2) **Stated objectives:**
To provide users with a rapid assessment tool to identify an organization's capacity in the area of Human Resource Management and to use these results to develop improvement strategies.

3) **Measurement focus:**
The HRM Assessment Instrument assesses civil service management. It is a self-evaluation tool designed to be used by NGOs and public sector health institutions to assess the status of Human Resource Management in their organizations and to facilitate the development of improvement strategies.

4) **Information generated**
The instrument is organized according to the core functions of a Human Resource Management system:
   • HRM Capacity
   • HRM Planning
   • Personnel Policy and Practice
   • HRM Data
   • Performance Management and Training

These functions are described in a matrix with four stages of indicators. For example, HRM Capacity:

	Stages of development for HRM and characteristics			
HRM Budget	There is no budget allocated for HRM staff or HRM activity within the organization.	There is limited money available to fund an HRM position or to conduct HRM activities...	Budget is allocated for HRM staff and related activities. Allocation is irregular and cannot be relied on for any useful long-range planning or the development of HRM systems.	Money for HRM staff and related activities is a permanent budget item, reviewed annually and adjusted if possible.
HRM Staff	There are no staff specially charged with responsibility for HRM functions.	There are HRM staff in the organization, but they have limited experience related to this field... and/or have other functions within the organization as well as HRM.	There are trained HRM staff in the organization, but only at a level to maintain basic procedures and record-keeping functions.	There are experienced HRM staff in the organization who maintain HRM functions. They participate in long-range planning for the organization.

5) **Methodology**

The HRM Assessment Instrument is a questionnaire that is answered according to which of four stages of development the organization has achieved. It is recommended that the instrument be used in conjunction with a review of documents at the organization including personnel files, job descriptions, financial/payroll records and relevant labor law. It is recommended that the HRM Assessment Instrument be administered by a committee of staff internal to the organization (e.g., Executive Director and representatives from Senior staff, Board of Directors and other staff members) with an external consultant facilitating the process. It can also be administered by an internal committee only. Together the group assesses each HRM component on the graph, agrees on the areas that need strengthening and formulates an action plan.

6) **Strengths & weaknesses**

Strengths:
- Broadly applicable tool for HRM self-assessment, despite rhetorical emphasis on NGOs and health organizations.
- Detailed response choices help reduce subjectivity in scoring.
- Extensive guidelines help organizations interpret the meaning of their responses.
- Limited resources required.

Weaknesses:
- Emphasis on framework and regulations rather than practice.
- Self-assessment, while important, leaves open the possibility of inflating results.
- Not pro-poor or gender-sensitive.

7) **Where to find it**

http://erc.msh.org/mainpage.cfm?file=7.40.htm&module=Toolkit&language=english

15. Analytical Framework for Institutional Assessment of Civil Service Systems, by Inter-American Development Bank

1) **History:**

Commissioned by the Inter-American Development Bank (IADB), the framework was developed to serve as a common reference to assess national civil service systems in countries located within the IADB's scope of action. It is based on World Bank, OECD and other reports. It was tested in the teaching and consultancy fields and published in 2002.

2) **Stated objectives:**

To facilitate the task of analysts of national civil service systems, providing a reference point and methodological directions useful for their work; to provide a common approach to diagnosis, facilitating comparative analysis of their conclusions; and to provide methodological measuring tools using indices that can allow comparison of certain aspects of different countries' situations.

3) **Measurement focus:**

The framework assesses **civil service management.** It is intended for any analysts of civil service systems, but especially those working for the IADB.

4) **Information generated:**

The framework divides HRM into seven interconnected subsystems, located vertically in three levels:
- In the upper level is HR Planning, which constitutes the entrance to any integrated HRM system, and allows the definition of coherent policies in all the remaining subsystems, to which it is connected.
- In the middle level there are five subsystems, organized logically in four, horizontal blocks:
    ◦ The Work Organization establishes and specifies the contents of the tasks and the characteristics of the people who will to carry out them.
    ◦ Employment Management includes the recruitment, movement and exit of the people.
    ◦ Results Management plans, stimulates and evaluates people's contribution.
    ◦ Compensation Management covers pay
    ◦ Development Management takes care of the training and development of individuals and the group.
- In the lower level, Human and Social Relations Management is connected to all the other subsystems.

Under the 7 subsystems are a series of critical points to be evaluated on a 5-point scale. Some examples are:

- *HR Planning follow-up allows flexible adaptation of personnel practices to the organization's strategic changes.*
- *Job descriptions are accurate, clearly explaining the structure of responsibilities. Each employee knows what he responsible for and how his contribution will be valued.*
- *Selection bodies are designed with professionalism and technical experience criteria, and act independently when performing their functions.*
- *People's performance is evaluated by the organization, comparing it to the expected standards.*
- *The salary structure is designed to attract, motivate and retain skilled people in the different positions required by the organization.*

One gender-sensitive critical point exists:
- *Mechanisms exist and are applied in recruitment and promotion procedures in the public employment to guarantee effective equality and nondiscrimination, to overcome disadvantages of sex, ethnic group, culture or origin.*

5) **Methodology:**

The framework comprises a report format, description of different components, and a checklist of indicators for evaluation. The final reports are intended to be a combination of quantitative results of the indices and subindices, plus qualitative analysis. Each implementer should adapt the method to the situation analyzed. Sources of information include desk research of standards, databases, external and internal research etc. and field research through interviews, panels of experts, and questionnaires.

6 **Strengths & weaknesses:**

Strengths:
- A broad approach, able to accommodate different perspectives and multidimensional approaches through adaptation.
- Well-suited for macro-analysis that characterizes institutional diagnosis.
- Adaptable for self- or external assessment.

Weaknesses:
- Method is presented in a compartmentalized manner that is difficult to follow and therefore not user-friendly.
- Indicators are not defined specifically, and supporting documentation may not be sufficient to ensure fair evaluation.
- Minimally gender-sensitive and not pro-poor.

7) **Where to find it**

http://www.iadb.org/IDBDocs.cfm?docnum=626933

16. Engendering Budgets: A Practitioners' Guide to Understanding and Implementing Gender-Responsive Budgets, by Commonwealth Secretariat

1) **History:**

There has been an explosion of gender-responsive budgets since 1995. However, basic information about gender budgets is not available for practitioners to understand some of the fundamentals and determine how best these could be adapted to their local context. This guide was published in 2003 to fill a specific gap in the literature and assemble in one document some of the background information that practitioners should have at their disposal to answer questions about how to design and implement a gender-responsive budget. It is based on analysis of past gender-responsive budgets and uses a similar framework.

2) **Stated objectives:**

To enable practitioners to design gender-responsive budgets adapted to their local conditions, as well as to assess the gender-responsiveness of existing budgets.

3) **Measurement focus:**

The analytical framework assesses **public financial management**. It can be used by policy-makers or by advocates for gender-responsiveness.

4) **Information generated:**
The analytical framework uses a five-step approach for analysis of existing budgets. A general description is given of each one along with suggestions of ways to evaluate it, but no explicit indicators are provided. The five-steps are:

1. Analysing the situation of women, men, girls and boys
2. Assessing the gender-responsiveness of policies
3. Assessing budget allocations
4. Monitoring spending and service delivery
5. Assessing outcomes

5) **Methodology: - type of instrument, methodology, data sources, respondents, participation in development**
The analytical framework is a descriptive guide. No explicit methodology is provided. Suggestions for data sources include national statistics, household surveys, local and international NGOs, international sources of on-line gender-specific information, national policy plans, and the national budget book.

6) **Strengths & Weaknesses**
Strengths:
- Comprehensive examination of how to ensure gender-sensitivity in budgeting based on wide experience.
- Flexible framework provides guidance without dictating what must be done, allowing for adaptation to local contexts.
- Pro-poor and gender-sensitive.
Weaknesses:
- Not a method but a guide – a genuine method should be produced before implementation is attempted.
- Method may not provide sufficient specific information for implementers unfamiliar with gender budgeting.

7) **Where to find it:**
The guide is available at http://www.thecommonwealth.org/shared_asp_files/uploadedfiles/%7BFBF59912-40C3-47A6-89C2-F3E5A0EA9B74%7D_Engendering%20Budgets%20final%20doc.pdf. A variety of supplementary materials is available at www.gender-budgets.org.

17. National Integrity Systems (NIS), by Transparency International

1) **History:**
The concept of the NIS was developed by Transparency International (TI) as part of TI's holistic approach to combating corruption. It was originally described in the TI Sourcebook and since been expanded and clarified in the NIS country study method. Since 2000, more than 70 NIS country studies have been carried out worldwide. In 2008 a new scoring method will be launched to complement the narrative country reports.

2) **Stated objectives:**
To assess the National Integrity System in theory and practice. The NIS encompasses the key institutions and sectors (the 'pillars') that contribute to integrity, transparency and accountability in a society.

3) **Measurement focus:**
The NIS method assesses **civil service management** and **leadership**. Given the breadth of topics covered in each NIS study as well as the range of cultural and policy contexts in which they might be produced, NIS studies can be used in a variety of ways. Primarily, they are used for advocacy efforts by non-governmental actors and by reformers in the government to help realize change.

4) **Information generated:**
The NIS is generally considered to comprise the following pillars: Executive, Legislature, Political Parties, Electoral Commission, Supreme Audit Institution, Judiciary, Civil Service/Public Sector Agencies, Law Enforcement Agencies, Public Contracting System, Ombudsman, Government Anti-corruption Agencies, Media, Civil Society, Business Sector, Regional and Local Government, and International Institutions. These 16 areas are covered in all country reports, with allowance for some local adaptation. The new scoring method covers all of these with the exception of Regional and Local Government.

Some examples of scoring questions are:
- To what extent are there provisions for a transparent selection process for civil servants?
- To what extent is citizen participation in civil service activities required by law?
- To what extent is there a process for members of the civil service to be held accountable for wrongdoing?

5) **Methodology:**
NIS country studies are narrative reports that will be paired with scores. Data are collected by a lead researcher who is generally an academic or other expert in the country under examination. All reports and scores are reviewed by the TI staff and by a validation committee in the country. One score is generated per NIS pillar by adding the scores of the underlying questions and transposing to a 1 to 7 scale. Researchers may adapt the method to account for local conditions; quality control for adaptation is done by TI.

6) **Strengths & Weaknesses:**
Strengths:
- Comprehensive look at a range of governance-related issues.
- Incorporates both hard data and expert evaluation for maximum topical coverage with minimized subjectivity.
- Produced in the country under examination by local experts, but with central oversight by TI.
- Method allows for adaptation to local circumstances, but standard components allow for comparison.
- Can be implemented with limited resources.

Weaknesses:
- Expert evaluation introduces some subjectivity.
- Broad method requires a range of expertise in order to be implemented.
- Not pro-poor or gender-sensitive.

7) **Where to find it:**
See http://www.transparency.org/policy_research/nis.

18. Diagnostic Framework for Revenue Administration, by World Bank

1) **History:**
As of the 1990s most support to modernization of tax and customs administration neglected certain important elements of the revenue administration system. As a result, many past operations may have failed to fully uncover the underlying causes of institutional and organizational dysfunction, leading to partial solutions and sub-optimal outcomes. The Diagnostic Framework for Revenue Administration was designed to fill this gap. Initial versions of the diagnostic framework were used in Latvia, Bolivia, and Colombia. It was developed in 2000 for a training course in China.

2) **Stated objectives:**
To provide a comprehensive diagnostic framework for revenue administration that takes into account, in an integrated manner, both the traditional areas of focus as well as areas that seem to have been under-emphasized.

3) **Measurement focus:**
The Diagnostic Framework for Revenue Administration assesses **public financial management**. It can be used to identify systemic deficiencies in the revenue administration system that lead to inefficiency and ineffectiveness in its operations.

4) **Information generated:**
The assessment is divided into eight frameworks for analysis of the revenue administration, each containing a detailed set of diagnostic questions: environment, resources; history; strategy; transformation processes; outputs; feedback mechanisms. Each question is accompanied by a possible reform option.

Some examples of the diagnostic questions are:
- Are revenue collection targets inflated due to political pressure to include additional expenditures in the budget?
- Does the number and type of taxes to be collected pose administrative difficulties for the revenue administration?
- Does the revenue administration employ effective strategies to monitor tax compliance and evasion?
- Are the outputs produced in a timely manner?

5) Methodology:
The Diagnostic Framework for Revenue Administration is an assessment questionnaire. There is no set of rules on how an assessment might be designed or implemented, or in what form the results should appear. The sole suggestion is that the framework can be used in a group setting involving different stakeholders, to guide brainstorming sessions and craft reform strategies.

6) **Strengths & Weaknesses:**
Strengths:
- Well-researched, extensive list of indicators for financial management assessment.
- Limited guidance allows for flexible implementation.
- Adaptable for self- or external assessment.

Weaknesses:
- Lack of guidance could result in less useful results among inexperienced implementers.
- Minimal gender-sensitive information.
- Despite goal of poverty-reduction, indicators are not pro-poor.

7) **Where to find it:**
http://siteresources.worldbank.org/INTTPA/Resources/DiagnosticFramework.pdf

INFORMATION SOURCES

1. **Country Policy and Institutional Assessment (CPIA), by World Bank**
   Criterion 13: Quality of Budgetary and Financial Management
   Criterion 15: Quality of Public Administration

1 **Stated objectives:**
To capture the quality of policies and institutional arrangements in International Development Association – eligible countries – the key elements that are within the country's control (not outcomes that are influenced by elements outside the country's control).

2) **Measurement theme and information generated:**
Criterion 13: Quality of Budgetary and Financial Management
This criterion provides indicators on **public financial management**. It assesses the extent to which there is: (a) a comprehensive and credible budget, linked to policy priorities; (b) effective financial management systems to ensure that the budget is implemented as intended in a controlled and predictable way; and (c) timely and accurate accounting and fiscal reporting, including timely and audited public accounts and effective arrangements for follow up.

Criterion 15: Quality of Public Administration
This criterion provides indicators on **civil service management**, **the policy-making system** and **service delivery**. It assesses the extent to which there is: (a) policy coordination and responsiveness; (b) service delivery and operational efficiency; (c) merit and ethics; (d) pay adequacy and management of the wage bill, along with justification for the rating. The score for the criterion is the average of these ratings.

Examples of indicators:
Criterion 13: Quality of Budgetary and Financial Management
- A rating of 3 indicates the following (according to the three dimensions listed above):
  (a) Policies or priorities are explicit, but are not linked to the budget. There is no forward look in the budget. The budget is formulated in consultation with spending ministries The budget classification system does not provide an adequate picture of general government activities. A significant amount of funds controlled by the executive is outside the budget (e.g., 10-25%), and a number of donor activities bypass the budget.

(b) Expenditures deviate from the amounts budgeted by more than 20 percent overall, or on many broad budget categories. Budget monitoring and control systems are inadequate. Payment arrears are 5-10% of total expenditures.

(c) Reconciliation of banking and fiscal records is undertaken less frequently than monthly, or discrepancies are not always accounted for. In-year budget reports are prepared quarterly less than 8 weeks after the end of the period, but their usefulness is undermined somewhat by inaccurate data or reporting only at high levels of aggregation. There are significant delays (e.g., more than 10 months) in the preparation of public accounts. Accounts are not audited in a timely and adequate way, and few if any actions are taken on budget reports and audit findings.

Criterion 15: Quality of Public Administration

• A rating of 3 indicates the following (according to the four dimensions listed above):

(a) Administrative structures are fragmented, and coordination mechanisms are generally inadequate to overcome parochial bureaucratic interests.

(b) Business processes can be overly complex, often causing unnecessary delays.

(c) Hiring and promotion formally merit-based, but there is extensive patronage in practice in several parts of government. Bribe seeking is accepted behavior in some agencies but not throughout government.

(d) Public employment as a share of total employment is higher than needed and unsustainable if adequate wages were paid. The wage bill represents an excessively large proportion of total government expenditure. Some sectors are overstaffed (particularly health and education). Pay and benefit levels are generally inadequate and there are major difficulties in attracting and retaining staff in key technical areas.

3) **Methodology:**

Scores are determined based on expert evaluation of a questionnaire. A separate and rating is assigned to each dimension under a criterion on a 6-point scale according to a detailed description of what each point means. Justification for each rating is also provided. The score for the criterion is the average of these ratings.

4) **Respondents:**

World Bank country teams

5) **Frequency:**

annually; latest release 2008

6) **Strengths & weaknesses:**

Strengths

• Extensive country coverage allows for wide comparison.
• Detailed ratings guidance minimizes subjectivity.
• Similar method reproduced over time allows for historical comparison.
• Simple, transparent method promotes understanding.

Weaknesses

• Evaluation based on judgments of experts external to the country open to criticism.
• Not developed in a participatory manner.
• Ratings are aggregates of diverse concepts and not all ratings are published, which renders them less useful for decisions on concrete action.
• While a separate CPIA criterion covers gender equality, criteria 13 and 15 are not gender-sensitive.
• While the ultimate goal of the CPIA is poverty reduction and a separate CPIA criterion addresses affects on the poor, criteria 13 and 15 are not pro-poor.

7) **Where to find it:**

http://go.worldbank.org/F5531ZQHT0

2. Governance Matters, by Government Effectiveness, World Bank

1) **Stated objectives:**
To evaluate governance, which is considered the fundamental basis for economic development, in 212 countries and territories.

2) **Measurement theme and information generated:**
The Government Effectiveness dimension provides indicators on **civil service management**, the **policy-making system** and **service delivery**. It measures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.

Examples of indicators:
- The Afrobarometer survey provides data from the following question: 'Based on your experiences, how easy or difficult is it to obtain household services (like piped water, electricity or telephone?' and 'Based on your experiences, how easy or difficult is it to obtain an identity document (such as birth certificate, driver's license or passport)?'
- Business Environment Risk Intelligence provides numbers from the Operation Risk Index on bureaucratic delays.
- The Bertelsmann Transformation Index provides numbers from its Management Index on consensus building, governance capability and resource efficiency.

3) **Methodology:**
Scores are determined based on aggregation of several hundred individual variables measuring perceptions of governance, drawn from 35 separate data sources constructed by 32 different organizations from around the world. A statistical methodology known as an unobserved components model is used to construct aggregate indicators from the individual measures. The aggregate indicators are weighted averages of the underlying data, with weights reflecting the precision of the individual data sources. Margins of error for each country are also calculated.

4) **Respondents:**
households, firms, businesspeople and experts

5) **Frequency:**
annually; latest release 2008

6) **Strengths & weaknesses:**
Strengths
- Complete country coverage allows for wide comparison.
- Subjectivity is limited by wide triangulation of data.
- Aggregation of many indicators reduces statistical error.
- Similar method reproduced over time allows for historical comparison.

Weaknesses
- Aggregate ratings are less useful for decisions on concrete action.
- Not developed in a participatory manner.
- Not gender-sensitive.
- Not pro-poor.
- Complex method and number of sources can inhibit understanding.

7) **Where to find it:**
http://info.worldbank.org/governance/wgi/index.asp

3. World Competitiveness Yearbook, by Government Efficiency, IMD

1) **Stated objectives:**
   To rank and analyze how the environment creates and sustains the competitiveness of enterprises in 55 countries.

2) **Measurement theme and information generated:**
   The Government Efficiency factor provides indicators on the **policy-making system** and **public financial management**. It includes criteria for public finance, fiscal policy, institutional framework, business legislation and societal framework.

   Examples of indicators:
   - Management of public finances over the next two years is likely to improve
   - General government expenditure as a percentage of GDP
   - Collected total tax revenues as a percentage of GDP
   - Government decisions are effectively implemented
   - The public service is independent from political interference
   - Public sector contracts are sufficiently open to foreign bidders

3) **Methodology:**
   Rankings are based on collection of both hard data (2/3 of total) and survey data (1/3 of total). Government efficiency is divided into 5 sub-factors composed of individual criteria. Each sub-factor has the same weight in the total. When Government Efficiency is aggregated with the other four factors, a ranking is determined.

4) **Respondents:**
   Leading scholars and IMD staff

5) **Frequency:**
   annually; latest release 2008

6) **Strengths & weaknesses:**
   Strengths
   - Emphasis on hard data reduces subjectivity.
   - Similar method reproduced over time allows for historical comparison.
   - Societal framework sub-factor includes criteria on equal opportunity, women in parliament, women in high-level positions and gender income ratio.
   - Societal framework sub-factor includes criteria on social cohesion as a government priority and income distribution.
   Weaknesses
   - Data not publicly available.
   - Ratings are aggregates of diverse concepts, which renders them less useful for decisions on concrete action.
   - Not developed in a participatory manner.

7) **Where to find it:**
   http://www.imd.ch/research/publications/wcy/index.cfm

4. Government at a Glance, by Organisation for Economic Co-operation and Development (OECD)

1) **Stated objectives:**
   To help OECD member countries to better assess, plan and measure their public sector reform agenda.

2) **Measurement theme and information generated:**
   The project will provide indicators on **civil service management**, the **policy-making system, public financial**

management and **service delivery.** It will encompass six categories of variables: revenues; inputs; public sector processes; outputs; outcomes; and antecedents or constraints that contextualise government efficiency and effectiveness.

Examples of indicators:
- General government revenue structure
- Compensation costs of employees as percentage of general government output
- Representation of women in senior positions
- Reduction and control of administrative burdens
- Trust in government
- Frequency of elections

3) **Methodology:**
Hard data and survey results will be collected by the OECD, the IMF, the UN etc. Datasets will be collected and presented without any overall score or measure.

4) **Respondents:**
OECD staff, OECD member states

5) **Frequency:**
first release projected late 2009, with regular updates anticipated

6) **Strengths & weaknesses:**
Strengths
- Designed for easy cross-country comparison.
- Data is presented without rankings, subjective interpretations etc.
- Sub-category on human resources management processes includes variables on participation of women in public employment.

Weaknesses
- Large number of data may be difficult to navigate for audiences without a specific interest.
- Not pro-poor.
- Not developed in a participatory manner.

7) **Where to find it:**
http://www.oecd.org/document/37/0,3343,en_2649_33735_37688524_1_1_1_1,00.html

5.   Bertelsmann Reform Index (BRI), by Management Index, Bertelsmann Stiftung

1) **Stated objectives:**
To assess 30 free market democracies on both individual policy fields and governmental performance.

2) **Measurement theme and information generated:**
The Management Index provides indicators on the **civil service management** and **leadership**. It evaluates a government's strategic capacity to introduce and carry out changes in a timely and resolute manner, in order to keep its democracy and market economy viable and effective. The first dimension of the Management Index, "executive capacity," analyzes a government's internal organization and procedures. The second dimension, "executive accountability," examines the extent to which certain actors or groups of actors with watchdog, informational and communicative functions in governance contribute to improving the knowledge base and the normative level of deliberation in policy-making.
Examples of indicators:
- How much influence does strategic planning have on government decision-making?

- Annual budget compared with the fiscal framework
- To what extent does the government consult with trade unions, employers' associations, leading business associations, religious communities, and social and environmental interest groups to support its policy?
- To what extent are citizens informed of government policy-making?
- Perception of one's own political influence

3) **Methodology:**

Scores will be calculated based on hard data, opinion surveys and expert evaluation of a questionnaire. Raw scores are translated to a 1 to 10 scale and then aggregated through an arithmetic mean. Scores are accompanied by country reports.

4) **Respondents:**

Three experts, including an economist and an in-country expert, evaluate each country.

5) **Frequency:**

First release projected for 2008, to be published biennially.

6) **Strengths & weaknesses:**

Strengths
- Incorporates both hard data and expert evaluation for maximum topical coverage with minimized subjectivity.
- Designed for easy cross-country comparison.
- Includes in-country experts to reduce criticism that data are generated "outside".

Weaknesses
- May replicate types of data that already exist.
- While a separate Status Index within the BRI covers some gender-sensitive indicators, the Management Index is not gender-sensitive.
- While the Status Index covers some pro-poor indicators, the Management Index is not pro-poor.
- Not developed in a participatory manner.

7) **Where to find it:**

http://www.bertelsmann-stiftung.de/cps/rde/xchg/SID-0A000F0A-CDD6FB92/bst_engl/hs.xsl/prj_52957_53423.htm

6. Open Budget Initiative, by International Budget Project

1) **Stated objectives:**

To provide citizens, legislators, and civil society advocates with the comprehensive and practical information needed to gauge a government's commitment to budget transparency and accountability in 59 countries.

2) **Measurement theme and information generated:**

The Open Budget Initiative provides indicators on **public financial management**. It groups questions into three sections:

Section One: The Availability of Budget Document
- Table 1. Budget Year of Documents Used in Completing the Questionnaire
- Table 2. Internet Links for Key Budget Documents
- Table 3. Distribution of Documents Related to the Executive's Proposal
- Table 4. Distribution of Enacted Budget and Other Reports

Section Two: The Executive's Budget Proposal
- Estimates for the Budget Year and Beyond
- Estimates for Years Prior to the Budget Year
- Comprehensiveness
- The Budget Narrative and Performance Monitoring
- Additional Key Information for Budget Analysis & Monitoring

Section Three: The Budget Process
- Executive's Formulation of the Budget
- Legislative Approval of the Budget
- Executive's Implementation of the Budget
- Executive's Year End Report and the Supreme Audit Institution

Examples of indicators:
- Does the executive's budget or any supporting budget documentation present information on policies (both proposals and existing commitments) in at least the budget year that are intended to benefit directly the country's most impoverished populations?
- Does the executive's budget or any supporting budget documentation present information for at least the budget year on extra-budgetary funds?
- Does the executive's budget or any supporting budget documentation explain how the proposed budget is linked to government's stated policy goals, by administrative unit (or functional category), for the budget year?

3) **Methodology:**
The Open Budget Initiative is based on a questionnaire consisting of 122 multiple-choice questions, plus four tables covering the manner in which budget documents are disseminated. The questions evaluate publicly available information issued by the central government. The majority of the questions ask about what occurs in practice, rather than about the requirements that may be in law. All of the questions were constructed with the intention that they should capture easily observable and replicable phenomena. Researchers and peer reviewers completing the questionnaires provide evidence for their responses. Sources include budget documents, laws or other public documents, public statements by government officials, and face-to-face interviews with government officials or other knowledgeable parties.

4) **Respondents:**
The researchers who have completed the Open Budget Questionnaire are from academic or other non-governmental organizations. International Budget Project staff undertook an analysis of each questionnaire, and in most cases spent several months in discussions with researchers to review the questionnaires. Following this review, the questionnaire was submitted to two anonymous peer reviewers.

5) **Frequency:**
Published 2006. No plans for repetition.

6) **Strengths & Weaknesses:**
Strengths:
- Well-researched and extensive survey of budget transparency based on good practice.
- Publicly available questionnaire could be completed by any interested organization.
- Simple methodology allows for easy implementation.
- Hard data minimizes subjectivity.
- Ample civil society involvement promotes broad-based support and attention to citizen concerns.
- Includes pro-poor indicators.
Weaknesses:
- Focus on transparency excludes other issues such as quality, effectiveness, etc.
- Not gender-sensitive.

7) **Where to find it:**
Homepage at http://www.openbudgetindex.org.

7. Cross-National Data on Government Employment & Wages, by World Bank

1) **Stated objectives:**
To provide systematic and comparable information on public sector employment and pay in more than 200 countries and territories.

2) **Measurement theme and information generated:**
The data provide indicators on civil service management. The Employment Data contain data on Total Public Employment, State-Owned Enterprise employees, and General Government, which is further divided into six mutually exclusive categories. The Wage Data contain data on Central Government Wage Bill, Average Government Wage, Average Government Wage to per capita GDP, Compression Ratio (the ratio of the highest salary to the lowest on the central government's main salary scale), Average Government to Manufacturing Wage, Average Government to Financial Sector Wage, and Average Government to Private Sector Wage.

3) **Methodology:**
Hard data presented in database form.

4) **Respondents:**
Compiled by World Bank staff.

5) **Frequency:**
produced irregularly; latest data 2000

6) **Strengths & Weaknesses:**
Strengths:
- Systematic and comparable information on public sector employment and pay worldwide.
- Unique breadth of information on employment levels and wage adequacy that can provide pointers for further analysis.
- Hard data minimizes subjectivity.

Weaknesses:
- Data of this kind at the international level can never be full comparable and must therefore be used with caution and in conjunction with other information.
- Accurate estimates of government employment are difficult to obtain and therefore some data may be questionable.
- Much of the data are outdated, depending on when country economists update it.

7) **Where to find it:**
http://go.worldbank.org/IJ9DCHUVK0

8. Integrity Indicators, by Global Integrity

1) **Stated objectives:**
To assess the existence, effectiveness, and citizen access to key national-level anti-corruption mechanisms across a diverse range of countries.

2) **Measurement theme and information generated:**
The Integrity Indicators provide information on **civil service management**, the **policy-making system, public financial management** and **leadership**. The indicators provide quantitative data and analysis of anti-corruption mechanisms, rule-of-law, and government accountability. In 2007, the Integrity Indicators were organized into six main governance categories and 23 sub-categories. Examples are:

Government Accountability
- Executive Accountability
- Legislative Accountability
- Judicial Accountability
- Budget Processes

Administration and Civil Service
- Civil Service Regulations
- Whistle-blowing Measures
- Procurement
- Privatization

Examples of indicator questions:
- Can the legislature provide input to the national budget?
- Can citizens access the national budgetary process?
- Are there national regulations for the civil service encompassing, at least, the managerial and professional staff?

3) **Methodology:**
Scores are calculated based on expert evaluation of a questionnaire. Each indicator score is then averaged within its parent subcategory, which produces a subcategory score. The subcategory score is in turn averaged with the other sub-category scores in a parent category. Category scores are averaged to produce a country score. Countries are grouped into five performance "tiers" according to their overall aggregated score: Very strong (90+), Strong (80+), Moderate (70+), Weak (60+), Very Weak (< 60). Scores are accompanied by country reports.

4) **Respondents:**
Each country is evaluated by a country team consisting of: a lead reporter who reports and writes the Reporter's Note-book, a lead social scientist who compiles the initial scores for the Integrity Indicators, at least three to five peer reviewers (a mix of other in-country experts and out-of-country experts) who blindly review the raw data and reporting.

5) **Frequency:**
First released in 2004, annually since 2006.

6) **Strengths & weaknesses:**
Strengths
- Comprehensive look at a range of governance-related issues.
- Narrative reports provide explanation for the issues behind the scores.
- Engages in-country experts so that data are generated from the "inside".
- Designed for easy cross-country comparison.
- Includes pro-poor indicators.
Weaknesses
- Expert evaluation introduces some subjectivity.
- Limited gender-sensitive indicators.
- Not developed in a participatory manner.

7) **Where to find it:**
http://report.globalintegrity.org/globalIndex.cfm