

Unclassified

GOV/PGC/GF(2009)2

Organisation de Coopération et de Développement Économiques  
Organisation for Economic Co-operation and Development

27-Apr-2009

English - Or. English

PUBLIC GOVERNANCE AND TERRITORIAL DEVELOPMENT DIRECTORATE  
PUBLIC GOVERNANCE COMMITTEE

Cancels & replaces the same document of 24 April 2009

## Global Forum on Public Governance

### Components of Integrity: Data and Benchmarks for Tracking Trends in Government

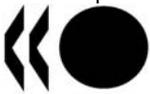
4-5 May 2009  
OECD Conference Centre  
Paris, France

*This document provides background material to support the discussion in Workshop 5 of the Global Forum on Public Governance. It was originally issued under GOV/PGC(2008)19 and approved by the Public Governance Committee under written procedure.*

For further information, please contact Janos Bertok. E-mail: [janos.bertok@oecd.org](mailto:janos.bertok@oecd.org), tel. +33 1 45 24 93 57.

JT03263695

Document complet disponible sur OLIS dans son format d'origine  
Complete document available on OLIS in its original format



GOV/PGC/GF(2009)2  
Unclassified

English - Or. English

## **COMPONENTS OF INTEGRITY: DATA AND BENCHMARKS FOR TRACKING TRENDS IN GOVERNMENT**

### **Introduction**

1. A key aim of the Public Governance Committee is to help governments better assess the implementation and impact of policies and measures in the public sector. This includes efforts to provide evidence-based comparative information on measures for fostering integrity and resistance to corruption in public organisations.

2. This document reviews strengths and weaknesses of existing methodologies for collecting information on corruption and integrity. It also presents an alternative approach and methodologies to collect data and develop benchmarks in key components of the integrity framework.

3. The draft report<sup>1</sup> was reviewed by the Expert Group on Conflict of Interest on 2-3 October 2008. The Expert Group agreed on the collection of data to be included in the “Government at a Glance” 2009 publication. The Public Governance Committee approved this report under written procedure in November 2008.

---

<sup>1</sup> The draft report GOV/PGC/(2008)19 was prepared by Professor Michael Johnston seconded to the OECD Public Governance and Territorial Development Directorate from Colgate University, Hamilton, New York, U.S, in close co-operation with János Bertók, OECD Secretariat.

**TABLE OF CONTENTS**

INTRODUCTION ..... 2

ACRONYMS..... 5

EXECUTIVE SUMMARY ..... 6

I. FROM MEASURING CORRUPTION TO ASSESSING INTEGRITY..... 9

    Data and benchmarks ..... 9

II. INTEGRITY: ISSUES OF THEORY AND MEASUREMENT ..... 12

    Complexities and caveats ..... 13

    Threats to integrity ..... 15

III. FIRST-GENERATION MEASUREMENTS: CORRUPTION ..... 17

    Why being cautious with corruption indices? ..... 17

    Notable efforts to measure corruption..... 18

    Perceptions and realities..... 20

    Methodological problems..... 21

    Practical difficulties ..... 24

IV. BUILDING AND SUSTAINING AN INTEGRITY FRAMEWORK ..... 27

    The Integrity Framework: Domains of assessment..... 28

    Providing evidence: Government at a Glance..... 31

    Sequencing: Implementing the strategy in phases ..... 31

    “Blinking indicators” and responses ..... 32

V. PHASE I: THE INTEGRITY FRAMEWORK ..... 33

    Key elements of the Integrity Framework..... 33

    Benchmarking the Integrity Framework ..... 37

    Government at a Glance: Data on core integrity instruments ..... 38

VI. PHASE II: RESISTANCE TO CORRUPTION AND RISKS: INPUT, PROCESS AND OUTPUT DATA ..... 39

    Basic examples..... 39

    Compared to what? Building benchmarks ..... 41

    Non-linear trends?..... 44

    Specific kinds of benchmarks ..... 46

VII. THE OUTER CONTEXT: OUTCOMES ..... 49

    Varieties of outcomes..... 50

    Sources of outcomes data..... 51

VIII. LINKING ASSESSMENT AND REFORM ..... 53

    Issue of validity, reliability, precision..... 53

    Implementing the DB strategy ..... 54

    Understanding the results..... 55

IX. GOOD GOVERNANCE AND INTEGRITY: DRAWING UPON THE INTEGRITY FRAMEWORK ..... 57

    A comprehensive approach..... 57

    Resource issues ..... 59

The OECD’s special role ..... 59

X. CASE STUDY: ASSESSING INTEGRITY IN PUBLIC PROCUREMENT ..... 60

XI. CONCLUSION: FEASIBILITY ISSUES..... 63

ANNEX 1. THE DATA AND BENCHMARK STRATEGY: POSSIBLE APPLICATIONS WITH  
RESPECT TO CORRUPTION ..... 66

**Tables**

Table 1. The Integrity Framework ..... 30

**Figures**

Figure 1. Trends in transaction costs as administrative steps proliferate..... 45

**Boxes**

Box 1. Measurement criteria: Validity, reliability and precision ..... 22

Box 2. Levels of measurement..... 23

**ACRONYMS**

CBSS	Council of the Baltic Sea States
CCBS	Crime and Corruption Business Surveys
CoE	Council of Europe
CPI	Corruption Perceptions Index
CSO	civil society organisation
DB	Data and Benchmarks
EITI	Extractive Industries Transparency Initiative
GaaG	Government at a Glance
ERS	Ethics Resource Center, United States
GRECO	Group of States Against Corruption
ICAC	Independent Commission Against Corruption
NGO	non-governmental organisation
OAS	Organisation of American States
PRS	Political Risk Services
TI	Transparency International
UNDP	United Nations Development Programme
UNODC	United National Office on Drugs and Crime
US	United States of America
WBI	World Bank Institute
WGI	Worldwide Governance Indicators

## EXECUTIVE SUMMARY

4. One of the most persistent problems for the past generation's anti-corruption reform movement has been a **lack of detailed measurements** and feedback. Corruption is largely a clandestine activity without immediate victims who might file reports; as a consequence, it is impossible to measure corruption directly with any validity, reliability, and precision.

5. Nonetheless, a **variety of corruption indices**, many based upon surveys of experts, businesspeople, and occasionally citizens, rank whole countries on a single dimension of "more" versus "less" corruption, or of more or less effectiveness at controlling corruption, for example. While such indices have been very valuable for raising awareness and keeping corruption issues in the public view, and are becoming considerably more diverse and sophisticated, the perceptions on which they draw are not the same thing as corruption itself, and the single scores they provide for whole countries provide virtually no detail on corruption as a day-to-day, sector-by-sector problem.

6. Such measurement problems have far-reaching consequences. Existing indices do a poor job at best of **tracking trends**, yet without such knowledge how to judge whether reform measures are having any effect. Corruption is a highly specialised and detailed activity, taking on many and varied forms and often existing in small niches of the public sector; while reforms must be equally detailed and adaptable, the best-known indices cannot provide such fine-grained knowledge and feedback. And even if the indices might point to places suffering the worst corruption, they give policy makers and managers limited guidance as to what to do about it.

7. What would be better?

8. This paper presents a Data-and-Benchmarks strategy that employs specific data on governance, not to measure corruption, but rather to facilitate broad-based assessments of **integrity**. The strategy begins with agency-specific data on public agencies' "Integrity Frameworks" – instruments and internal systems of incentives, controls, and goal fulfilment – which are nested within broader assessments of resistance to corruption and risks to integrity, and finally within comprehensive assessments of context, including outcomes.

9. At the level of agencies' own **integrity frameworks**, assessments will primarily focus on the "hardware" of integrity management: systems of defining integrity standards, guidance, monitoring compliance, control, reward, and sanctions, moreover, on relevant conditions in the wider public management and governance context such as provisions for transparency, accountability and citizen input. The approach proposed by the Data-and-Benchmarks strategy is to move from a static description of components of the "hardware" of the Integrity Framework to data and benchmarking capturing the dynamic evolution of formal integrity instruments. In addition, detailed surveys and interviews with public officials may provide complementary information in the form of factual knowledge on functioning and impact of integrity instruments – for example awareness and commitment of staff or how often they witness acts of official misconduct – rather than asking for impressionistic ratings.

10. **Analysis of risks to integrity**, among a variety of vulnerabilities to integrity including corruption, constitutes the starting block of the Integrity Framework. Resistance to corruption and risks is proposed to be assessed through data on government activities – for example, degree of using exceptions to competitive tendering modification in contract management and how quickly and accurately invoices are paid; how much time and how many steps are involved in obtaining a license. Those data can be compared to similar evidence gathered in comparable agencies, and to benchmarks embodying the norms of performance across those governmental entities. If issuing a license in city A involves many more steps and much delay when compared to other cities, for example, likely seeing

the **incentives for corruption** in public organisations (officials have learned ways of extracting income through added requirements, “lost” documentation, delays and the like) and for businesses and citizens (frustrated by long delays and pervasive uncertainty, will be more tempted to bring the process to an end).

11. As examples in the paper show, such data can be very detailed, and can identify agencies or governments where problems are pronounced. The data can track trends over time, enabling policy makers and managers both to judge the effects of reforms in considerable detail and to reward governments, officials, and communities for positive change. Most important, the results of such assessments of corruption risks and vulnerability are **actionable**, pointing directly toward steps that need to be taken: where slow processes are perpetuating corruption officials know where they must speed things up. The proposed data and benchmarks will emphasise a series of “blinking indicators” that reveal particularly serious difficulties and top priorities for reform.

12. Finally, the strategy emphasises not just “outputs” or “performance” of governments, but also envisage moving towards **outcomes**. Do officials know, and respond effectively to, social problems and the wishes of citizens? Are resources used not just honestly and efficiently, but also effectively, for the wellbeing of communities and societies? Do citizens perceive their government agencies and officials as honest and responsive, and are they confident that they can voice their wishes and complaints in ways that are taken seriously by those who make and implement policy?

13. In the long run, data and benchmarks in these domains, and comparisons among them, are essential to assessing levels of integrity. In all these domains – integrity frameworks, corruption risks and vulnerability, and outcomes – the strategy is to assemble valid, reliable **data** – instead of relying upon perceptions – and to compare them with **benchmarks** compiled across comparable government organisations. Over time such data will allow policy makers and managers to reach comprehensive assessments of integrity processes and problems as dynamic influences upon governance on a very broad scale, while also identifying specific, tightly-targeted measures for improvement – measures whose effects can be judged with a high level of confidence.

14. It is important to understand that the DB assessment strategy presented in this paper is not another corruption index. Integrity is conceived of in terms of positive values; at the other end of the continuum not corruption but rather vulnerabilities of several sorts can be found. Key comparisons will be made, not nation to nation, but rather at the levels of programmes and public organisations, over time; the key is not to rank them in terms of integrity but rather to provide detailed, relevant, actionable knowledge that managers and political leaders can use to address integrity problems with precision and confidence. There are several sound incentives – not least, the opportunity to take credit for positive change – that will encourage managers, political leaders, and public organisations to take part in the DB strategy for assessing integrity.

15. Implementing such a detailed yet broad-based strategy will involve careful efforts to assure officials and governments to co-operate, provide data, and discuss results. A key point in that process of persuasion is that the results of the data and benchmarks process will be **useful** to agency managers (*e.g.* providing relevant warning signals on risks) and political leaders, not threatening to them. Further, participation need not be resource-intensive: many of the sorts of data involved are items that governments gather anyway, or may easily compile with small administrative changes. The data will allow them to demonstrate progress toward major positive goals of good governance – not require them to engage in a process of labelling themselves, or their governments, as more or less “corrupt” – and, equally important, to take credit for success. That last point is a critical, yet often-overlooked, aspect of official commitment toward integrity and public sector reform.

16. The last chapter present a case to illustrate how a data-and-benchmarks process might be initiated in a key area of government – that of **public procurement**. Proposed data sets highlighting the vulnerabilities in public procurement will be included in the OECD flagship publication

*Government at a Glance*. The proposed phased process data-gathering and benchmarking can be launched in non-disruptive ways, shaped by learning to yield results useful at many levels.

17. The overall goal is to emphasise positive values of integrity, justice, transparency, accountability and efficiency, and to demonstrate and reward progress on those dimensions, rather than to stigmatise public officials, agencies, or whole communities as “corrupt”. The strategy as envisioned here will dovetail with the OECD’s larger *Government at a Glance* project, its continuing efforts to enhance good governance, its longstanding competence in assembling and analysing economic and public-policy data, and with its emphasis, in the *OECD Integrity Framework*, on positive values in public service, government and society. When fully operational it will give the OECD a unique scope and competence in the assessment of integrity and resistance to corruption as key pillar for achieving good governance.

## COMPONENTS OF INTEGRITY: DATA AND BENCHMARKS FOR TRACKING TRENDS IN GOVERNMENT

### I. FROM MEASURING CORRUPTION TO ASSESSING INTEGRITY

18. The OECD has unique opportunities to shape the policy debate as a major actor in the international movement in the field of integrity-building and assessment. Efforts over the past decade to measure corruption and “good governance” and to rank whole countries in terms of perception data have proven useful in terms of raising awareness. But when it comes to targeting resources upon specific reforms, and in particular to assessing progress over time, far more detailed evidence and a more comprehensive view of what is desirable in government are needed.

19. Via its conceptual frameworks, such as the Ethics Infrastructure concept and the Integrity Assessment Framework, the OECD possesses a sound fundamental approach. It also has significant resources: a generation’s worth of anti-corruption expertise; unparalleled networks, information and data resources; and strong credibility based both on anti-corruption and integrity instruments, such as the 1997 *OECD Convention on Bribery of Foreign Public Officials in International Business Transactions*, 1998 *OECD Recommendation on Improving Ethical Conduct in the Public Service* and the 2003 *Recommendation on Guidelines for Managing Conflict of Interest in the Public Service*. All grow out of a half-century of effective dealings with the world’s major economies. More recent analytical efforts, such as the *Principles and Checklist for Enhancing Integrity in Public Procurement* and the forthcoming *Government at a Glance* series, are further evidence of the OECD’s contribution in these fields.

20. But the OECD also faces significant challenges with respect to integrity assessment, some in fact growing out of the leadership role outlined above. Developing relevant data and measurements that are both more detailed and part of a comprehensive view of integrity is in many respects a venture into unknown territory. Data and analysis must be detailed but not burdensome; the resulting measures must be of genuine use to governments and public officials. The assessment strategy must offer real incentives for co-operation, mesh well with existing data and analysis processes, and point directly to appropriate reforms. Above all it must yield results that are easily and clearly understood, and that draw attention to successes.

#### **Data and benchmarks**

21. This paper proposes a twofold strategy – gathering **data** on governments and their activities, and assessing those data against detailed and broadly-comparable **benchmarks**. Rather than attempting to estimate overall levels of corruption as a free-standing problem, the strategy to be outlined here begins with corruption vulnerabilities and corruption resistance (Gorta, 2006) and links those concerns both to public agencies’ internal climates and to broader outcomes in – and participation by – society as a whole. The key throughout is the assessment and pursuit, not of corruption as such, but rather of **integrity** – thought of in terms of justice, transparency, accountability and efficiency – providing detailed, valid information and feedback useful to those charged with pursuing the positive values it embodies. This proposed Data and Benchmarks (hereafter, “DB”) strategy is not without its gaps and difficulties. It requires considerable co-operation and lasting commitments to honest implementation. And it will not be appropriate or feasible everywhere. Still, it represents a step forward that fits well with both the current state of integrity-building efforts and the OECD’s established strengths.

22. Most of the data involved are objective, in the sense that they do not rely upon perceptions reported by distant observers. Others are subjective but draw upon officials’ and citizens’ actual experience; they bring a wide range of interested parties into the evaluation process, building both accountability and credibility. Targeted assessments of where a specific agency or programme stands in terms of controlling corruption, vulnerabilities and indicating specific countermeasures can be made, and then can be used for subsequent assessments of progress. Factors building or increasing resistance to corruption can be identified, and linked not just to quantifiable outputs of government but also to more general **outcomes** – a key goal of the *Integrity Assessment Framework* developed by the OECD (OECD 2005).

23. The DB strategy will enable successful leaders, policy makers and managers to learn from documented experience and take credit for their efforts while allowing citizens to see actual results of their participation. Conversely, leaders and officials who fail can be identified and, potentially, held accountable. All of these aspects of integrity can be assessed in terms of real, easily-understood, units of measurement – not just arbitrarily defined points on a perception scale. Conceivably most important, **trends over time** can be tracked in detail – an essential element of any integrity-building effort, and of figuring out “what works”, yet one that has proven elusive for first-generation efforts focusing on comparing levels of corruption.

***What the data and benchmark strategy is – and is not***

24. It is important to be clear about what the DB strategy is, and is not. It is not another corruption index, nor is it aimed at country-by-country rankings. It is, instead, an attempt to assess levels of integrity, defined in terms of positive values and accomplishments, and a range of vulnerabilities – in effect, the consequences of past problems and a range of current incentives and opportunities – that undermine the responsiveness and quality of government and public services. Comparisons will be made at the organisation and programme levels, *not* in order to create integrity or corruption rankings among them, but rather to track changes and developments, positive as well as negative, in programmes and organisations over time.

25. One way to understand these points is to think of a continuum. At one end is integrity, embodying key values of justice and equity to promote the public interest; transparency and openness; accountability; and efficiency.<sup>2</sup>



26. At the other end of the continuum is, not corruption, but rather a number of vulnerabilities: risks of error, poor responsiveness, insufficient supervision and accountability, uncontrolled discretion, and a wide range of resource problems – vulnerabilities that include, but are by no means synonymous with, risks of corruption. While, by definition, the absence of integrity cannot be measured, the data and benchmarks this paper discusses can give policy makers and managers sound estimates of the scale and locations of those kinds of vulnerabilities. The data and benchmarks proposed can also suggest actions to

---

<sup>2</sup> ‘Efficiency’ is deliberately ranked fourth among those values, and must be understood and assessed in more selective and limited ways than we might apply it to the private sector. Many public functions, such as education, the justice system, and corrections, do not produce a discrete “output”, others, such as many public transport services, are parts of the public sector rather than the private because their provision is inherently inefficient; and still others, such as sanitation or environmental protection, yield public goods that are unrelated to “efficiency” as such.

be taken to address the vulnerabilities and can give policy makers and managers a sound sense of whether or not progress is made in dealing with them.

27. Why would an organisation and its managers wish to provide and analyse data, and why would they agree to specific benchmarks? Because it does not ‘measure corruption’ and because the organisations and managers to be included in the process would participate in identifying key values and in designing applications of the DB strategy appropriate for their own organisations, several **positive reasons** to participate become clear:

- **Personal** incentives: managers and political leaders can claim credit for improvements and successful governance and management.
- **Professional** incentives: professional managers and service deliverers can increase their standing among colleagues with positive results or changes demonstrated by DB processes
- **Budgetary** incentives: ideally, high-integrity organisations and those moving in positive directions could be provided with additional resources, and as a practical matter will be in a far stronger position to resist allegations of fraud, waste, abuse, and corruption, and to argue against budget cuts on the basis of such concerns.
- **Cultural** and national incentives: countries suffering from the ‘labelling’ problems created by corruption indices will be able to demonstrate to critics, potential investors, international backers, and their own citizens that governance is improving; and finally
- Incentives in the form of heightened **public credibility** for agencies, and for government in general – a development that can relieve a certain amount of pressure on public budgets over the long term.

28. Indeed, over time problems that appear to be corruption-related, or that are alleged to be by hostile critics, may well have other origins, a possibility that can be demonstrated through careful and repeated gathering of data and benchmarks. Where that is the case a considerable amount of tension and uncertainty connected with oversight processes may be alleviated, and organisations can make a persuasive case that additional resources sent their way will be put to effective use.

29. None of these sorts of concerns can be addressed by using corruption indices – either those that are now in hand, or any improved corruption measure we could conceive of – for, to reiterate the key point made above, *integrity consists of positive values*, and is not the obverse of corruption. For that reason, the data-and-benchmarks strategy proposed in this paper is a fundamentally different exercise.

## II. INTEGRITY: ISSUES OF THEORY AND MEASUREMENT

30. Assessments of integrity will encounter significant methodological challenges. As with corruption indices, whose virtues and problems will be discussed below, any discussion of integrity raises questions about definitions, and raises important normative issues that must be acknowledged. But even more basic, as Christiane Arndt (2008; see also Arndt and Oman, 2006) argues, is the need for **theory**:

- What should governments do, how should they do it, and where are the main determinants of success or failure to be found?
- Who decides, on what criteria, whether success or progress is being attained?
- How do the various components of integrity relate to each other, and is there a hierarchy among them?

31. These theoretical questions, and the normative values and preferences that animate them, have been debated since the time of the Greeks (Dobel, 1978), and the role of *virtu* in the arts of government have been an issue since at least the time of Machiavelli (see, for example, Shumer, 1979). Such questions can scarcely be resolved here; indeed, neither integrity, corruption, nor the proper role of government will ever be defined in a neat formulation.

32. A sound starting point is offered, however, by Jocelyne Bourgon's (2007) idea that "Public administrations are a vehicle for expressing the values and preferences of citizens, communities, and society as a whole."<sup>3</sup> Officials and agencies who understand those values and preferences, pursue them in good faith, and are open to scrutiny and feedback from "citizens, communities, and society as a whole" are governing with a high level of integrity. However, integrity is more than the absence of corruption, or of other sorts of abuses; a working definition must rest upon important positive values and include the **outcomes**, not just outputs, of government. Van der Wal, Huberts, van den Heuvel, and Kolthoff (2006), comparing public- and private-sector ethics, add that:

33. Integrity, meaning "wholeness" in its original sense (the Latin word *integritas*), is...(acting or being in) accordance with the moral values, norms and rules, valid within the context in which one operates... Within an organisational context, it denotes 'the quality of acting in accordance with the moral values, norms and rules accepted by the members of the organisation and its stakeholders. Integrity is a quality or characteristic of individual or organisational behaviour; it may even be considered an individual and organisational virtue.

34. This paper pushes such notions further, taking "wholeness" to include not only the behaviour of officials and their organisations but also the internal climate of public institutions – their stated procedures, informal norms, divisions of labour, incentive and accountability systems, monitoring processes, and use of resources, to name some major aspects – *and* to interactions with the wider society. Assessing integrity in public institutions, as opposed to businesses and other private-sector bodies, requires that breadth of focus – an issue to be discussed below.

---

<sup>3</sup> To be sure, governments taken as a whole include more than just the administrative apparatus, but for reasons to be outlined below the DB strategy to be proposed here will deal primarily with the administrative core and public service, rather than with representative or judicial institutions.

35. Like the “wholeness” that is the main focus for Huberts *et. al.*, those broader aspects of integrity are expressions of, and can be judged against, the basic values of society. Not surprisingly for such basic normative issues, just what those values *are* is a complicated issue; van der Wal and colleagues (2006: 342) identify many possibilities for the public and private sectors, and even after a rigorous effort at boiling them down to essentials still arrive at no fewer than twenty-one core values for organisations, arrayed on a broad public-private continuum. Moreover, precise lists of values, and their social significance in the broadest sense of that term, will differ from time to time and place to place.

36. Thus **core values of integrity** must be specified in terms of a few broadly applicable values. Considering this the followings list core values that are interdependent:

- **Justice, equity to promote the public interest**
- **Transparency, openness**
- **Accountability**
- **Efficiency**

37. Any list of core values of integrity is rightly open to challenge and debate. Moreover, none of the terms in the list is easily defined. In this paper, however, **justice** refers to actions and decisions by officials, agencies, and society that reflect and enhance basic fairness, equity and respect for human rights and dignity in order to promote the public interest.

38. **Transparency** empowers citizens, journalists, civil society groups, and other outsiders to examine the operations of a public organisation, to publicise their findings and act upon them, and, in some instances, to intervene in official dealings where transparency has been restricted or problems have been revealed. It imposes reciprocal obligations upon officials and agencies to make information available, via procedures and within limits that may vary significantly from one agency and society to the next, and to provide channels for public feedback. Whether or not that feedback is taken seriously is an important issue, but one that falls generally under the heading of accountability.

39. **Accountability** is the ability of those who delegate power and resources to require the officials and institutions entrusted with them to explain, justify and document their actions, and to demand that substantive and procedural shortcomings be addressed. Accountability has internal dimensions (principals, agents, and clients of agencies must be accountable, in varying ways, to each other) as well as external ones (public agencies and officials must be answerable to citizens and, again in somewhat different ways, *vice versa*).

40. **Efficiency**, in a strict sense, may be thought of as maximising units of output (say, miles and quality of paved roads) per units of input (funds, labour, materials), much as in private business. Clearly, a public organisation that wastes money, pursues self-serving goals while ignoring legitimate citizen demands, and produces poor-quality services and facilities is not contributing to overall integrity (and is likely to offer numerous opportunities for corruption). But “efficiency” in the public sector is a complex and problematical notion, and assessments must be made with great caution, as will be discussed below.

### **Complexities and caveats**

41. All four of the values above are laudable goals. All are also open to dispute in terms of their practical, day-to-day meanings. None of them is simple to measure or assess.

42. **Justice** is often a matter of opinion. Because of the government’s role in dealing with, responding to, and at times resolving conflicts among competing groups and interests in society, what constitutes justice for one community may be a palpable injustice in the eyes of another. Further, justice requires legitimacy if the authoritative allocation of values, and the actual or threatened uses of force, that may be involved in governing are to be acceptable rather than constitute new grievances in their own right. Justice may be a procedural matter in some cases: a party to a dispute that has been accorded due process and fair treatment may have been treated justly and impartially even when decisions go against their wishes. In other instances, though, justice is a substantive issue, and in those cases differing views and interests may be difficult to reconcile. Honest differences over what constitutes justice are not just possible within government but are among the primary forces driving political processes: deep differences of opinion may exist as to what constitutes a just apportionment of tax burdens, over the limits of tolerable dissent, the type and extent of punishment appropriate to a given crime, or regarding the question of whether public support for the fine arts is a luxury or an essential affirmation of a community’s culture and identity. Integrity assessments cannot and should not attempt to resolve such questions (indeed they should be “policy-neutral”, as will be discussed below). More often they will focus upon the place and strength of justice within an agency’s internal climate of values and accountability to promote the public interest.

43. **Transparency** is often used as a catch-all term for integrity. But it means little if the goals government pursues are unjust, and if it is not accountable. When it comes to conventional notions of efficiency, transparency measures may exact significant costs. Transparency as a core value of integrity is thus contingent upon whether and how justice, accountability, and efficiency are upheld as values in their own right. At the same time, however, official promises to honour those three values mean little if they cannot be verified and are not subject to feedback. Transparency therefore has a definite place in the list of core values of integrity, both as a way of assessing commitment to the other values and as a critical factor enabling the data-and-benchmark activities to be outlined below.

44. **Accountability** is difficult to maintain in any situation, but particularly so in government. Public agencies are large, complex, and (as will be noted in the discussion of efficiency, below) shaped by influences that differ significantly from those found in the private sector. The extent to which public purposes are being accomplished, and public power and resources used effectively, may be a matter of considerable dispute; governments cannot point to a “bottom line” or to share prices as proof of success. Ideally public officials are accountable to, and both driven and constrained by, many segments of society. In democratic societies political contention revolves around contrasting views of the role and goals of government – views that do not go away between elections. Many of the costs of public functions are socialised across society as a whole, including many who are not direct beneficiaries or “users”. Indirect beneficiaries of a service (such as education, for those who do not have children in public schools) might not perceive their own stake in those services; some of those indirect beneficiaries might see most or all expenditures in such areas as wasted. The situation is quite different from that of the private sector, where purchasers or consumers of most goods enter into (or opt out of) transactions by choice, may have several possible suppliers to deal with, enjoy personal benefits from their purchases, pay all or most of the cost of the goods plus a premium representing profit for the provider, and can (at least in theory) take on the private provider on an equal legal footing in the event of disputes. Under those conditions accountability is a much more straightforward concept.

45. **Efficiency** would seem to be the simplest of the four core values, but is no less complex. Much of what governments produce is intangible (recognition and other symbolic benefits, protection of rights, and due process). Other “outputs” – clean air and water, for example, or conflict resolution – are public goods that, unlike street paving, are difficult to measure in monetary terms, require co-operation by private parties, and in key respects are impossible to link to specific inputs. Public-sector “efficiency” can be a quite different thing from the quality of services: it is easy to imagine an agency that executes flawed (or counterproductive) policy with consummate efficiency. Most public agencies do not function as market

entities; policing, for example, could not work if it did not rest upon Weber's state "monopoly of the legitimate use of physical force" (Weber, 1946 ed.: 78), delegated, to be sure, among levels and areas of jurisdiction. But that same monopoly raises complex issues of rights, due process, and restraint that are both outputs and outcomes even though they cannot be quantified. Monopoly status also affects functions like procurement; for many commodities (computers or paper clips) government may be the largest single consumer, or nearly so, and might be expected to exercise market power other purchasers do not possess. In other areas – the procurement of aircraft carriers, for example – a government may be the only "customer" and suppliers might be few in number as well. Again, conventional notions of efficiency, such as economies of scale, may not apply. Finally, efficiency in the public sector may include the speed and accuracy with which government responds to changes in public preferences, or to emerging problems and circumstances (an accountability issue too). While it might be said that private business must also respond to changes in mass preferences, they do so in a setting of clearer signals (via the market), more options (including the option *not* to respond), and fewer constraints. Thus any assessment of public-sector efficiency, and the choices as to how to weigh it as a component of integrity, must be based on an understanding of the distinctive realities of the public sector.

### Threats to integrity

46. Just as integrity itself is a complex concept, so too is the environment within which it must be pursued, defended, and assessed. Threats to integrity come in many forms and from numerous sources. Corruption is perhaps the most commonly-emphasised threat, for understandable reasons, and the renewed emphasis and study it has received in recent years have been welcome developments. But other threats concern us as well:

- **Weak or abusive internal accountability:** officials may be poorly supervised and evaluated; or, accountability links and processes may be used to encourage, require, conceal, and reward inappropriate activities.
- **Weak or compromised external accountability:** feedback from the public and other outcome assessments may be insufficient, biased, or devalued; public participation may be discouraged or used to conceal – or justify – poor outcomes; agencies and officials may be unresponsive to citizens and to goals of policy.
- **Political interference:** while political officials can and should exercise oversight over public agencies, such interventions can become self-serving, abusive, or disruptive. Distinguishing between legitimate oversight and political interference may not be easy. While excessive intrusion of politics into administration is a clear threat to integrity, efforts to isolate the two from each other can end up producing an unaccountable administrative apparatus – one that may well serve its own goals first.
- **Poor internal administrative controls:** many familiar vulnerabilities fall under this category, including insufficient control over funds, supplies, and information; poor recruitment and training practices, theft of time and services, and so forth.
- **Unclear/contradictory/unattainable/illegitimate goals:** officials who are not sure of the purposes served by their jobs and agencies, or who are expected to pursue goals that are too numerous, contradictory, cannot be attained, or are illegal or otherwise undesirable are unlikely attain high levels of integrity.
- **Resource problems:** agencies that are starved of key funds and resources, or that are given the wrong mix of resources, will serve few positive values and be unable to respond to the public or

stated policies. But too many resources can encourage waste or diversion of goods to private purposes, while procurement from the wrong sources can be a sign of waste, corruption, and political interference.

- **Faulty division of labour:** who is supposed to do what in an agency, answerable to whom, is a key question. Where tasks are overlooked or put into too many hands, key values will go unaddressed. Of particular concern in this regard are **monopolies** – often, a risk of corruption and weak accountability.
- **Inappropriate or unchecked discretion:** no set of working rules can anticipate all contingencies, and a degree of discretion is usually essential. Appropriate levels of discretion can keep morale high and the connection between means and ends clear. But uncontrolled discretion invites abuse, waste of resources, inconsistent outcomes, weak accountability, and external interference. Eliminating discretion can also do harm, however, by discouraging innovation, reducing morale, and impeding responsiveness to the public as officials fall back upon “by-the-book” conduct.
- **Inappropriate social expectations:** while agencies and officials must listen to the views of citizens, the latter should be aware of what can and cannot be expected of government, and those standards must be followed consistently. Officials must be able to distinguish appropriate from inappropriate public expectations and be prepared to distinguish between legitimate innovation and the abuse of power in responding to new and unusual demands that may fall into grey areas.
- **Clientelism:** too close a relationship between political leaders and officials, or between them and citizens, can produce a situation in which power and resources are used less to pursue stated policy goals than to reward personal or political allies and clients. Favouritism is a closely-related threat as well; in both cases access to services and benefits becomes less a matter of legitimate qualification and need than a question of “who you know” and, often, what one is willing to do for the politician or leader providing the benefits.
- **Capture:** just as political and bureaucratic influence can extend too far into society, private interests can intrude upon government functions. In many cases the result is corruption of familiar forms, but in some instances the result is the colonisation or outright capture of agencies by private interests. Regulatory agencies overseeing a particular industry may be “captured” by that industry, through a number of types of connections. A related risk is “soft capture”, in which an agency becomes so dependent upon an industry for information or implementation of policy that it loses the ability to act on its own.

### III. FIRST-GENERATION MEASUREMENTS: CORRUPTION

47. While integrity is a complex notion, and threats to it are diverse, most assessment efforts over the past two decades have been devoted to the measurement of corruption (for discussion and evaluations see Arndt and Oman, 2006: Ch. 1, 2; Arndt, 2007; Cameron, Chaudhuri, Erkel, and Gangadharan, 2005; Kenny, 2006; Kurtz and Schrank, 2007; Søreide, 2005; Thomas, 2007; on East Europe and Central Asia, see Knack, 2006; on Latin America, see Transparency International, 2006; see also Doig and McIvor, 2003; and the United Nations Development Program Governance Indicators Project<sup>4</sup>). Most such efforts produce whole-country corruption scores based entirely, or primarily, upon perceptions data, although intriguing variations on that approach exist as well. Both the indices produced and the many debates they have touched off are of indisputable value – not least for helping direct attention to regimes where governance is in dire need of improvement.

48. But a sustainable integrity framework requires more policy relevant, detailed and understandable data. Perception-based national indices do not help policy makers and managers understand or attack corruption problems in focused ways. They are even less useful for assessing progress (an excellent review of such issues appears in Sampford, Shacklock, Connors, and Galtung, 2006). If, as is often proposed, policymakers link aid and other incentives to improvements in the quality of governance (see, for example, MCC, 2007; MCC, 2008; for a more general discussion, Søreide, 2005) policy makers and managers will need evidence that enjoys solid confidence on all sides. Less frequently noted is the fact that recent international anti-corruption agreements – beginning with the OECD Convention on Combating Bribery of Foreign Public Officials in International Business Transactions, but also including the GRECO process in the Council of Europe, the OAS Inter American Convention Against Corruption, and the United Nations Convention against Corruption – have been, or will soon be, generating data in unprecedented volume and detail. Much the same is true of e-governance initiatives. In some respects the time has never been better to embark on new assessment methods and to seek new forms of co-operation based on reviewed experience and lessons learned in assessment.

#### **Why being cautious with corruption indices?**

49. Corruption indices have become so numerous and diverse that the United Nations Development Programme (UNDP) and the NGO Global Integrity recently published a report intended to help anti-corruption advocates sort through the alternative methods and results they offer. Indeed, an appendix to the report lists fifteen major efforts to measure corruption on a global or regional basis (UNDP and Global Integrity, 2008). Such indices have become increasingly sophisticated, yet significant problems remain. One basic issue is definitions: consensus on a nominal definition of corruption likely will never be reached, and measures based on contested – or merely vague – concepts are inherently problematical (Brown, 2006; Johnston, 2005a; Kurer, 2005). For analysis at a general level it is tempting to focus on unambiguous core cases and not worry about the boundaries of the concept – one reason, perhaps, why many models of corruption, like many of the surveys that go into calculating corruption indices, effectively treat “corruption” and “bribery” as synonyms. But when it comes to measurement and changes in real societies the margins become critical.

50. Further, corruption is a hidden phenomenon: all who know of a corrupt act usually have an interest in concealing it and, unlike many forms of conventional crime, there is frequently no immediate victim with a reason to file a report. Where corruption is most serious many officials charged with control

<sup>4</sup> Background documents and details at [www.undp.org/oslocentre/flagship/governance\\_indicators\\_project.html](http://www.undp.org/oslocentre/flagship/governance_indicators_project.html) (Viewed 5 September 2008).

are themselves compromised, and reporting corruption can be an exercise in risk and futility. Statistics on conventional crimes are notoriously inaccurate; how can analysts measure an activity that is usually clandestine?

51. Then there is the problem of evidence: what should be counted or measured in practice?<sup>5</sup> Most indices are based, to varying degrees, upon **perceptions**. Perceptions are not the same thing as corruption itself; indeed, they can be erroneous, misleading, or self-serving, and are vulnerable to “echo chamber” dynamics if people asked to assess levels of corruption end up repeating “what everybody knows”. Add to this the complex relationship between corruption and scandal (Moodie, 1980; Markovits and Silverstein, 1988) – either may be found in the absence of the other, and scandal in particular is frequently used as a political weapon – and it becomes clear that perceptions and corruption may be very different things. Further, it is hard to say what perceptions of “a high level of corruption” actually mean – numerous cases, large monetary stakes, corruption at high levels, cases attracting major public attention, or simply corruption that takes place out in the open (Rose-Ackerman, 1999)? Many strategies commonly used to deter or reveal corruption – bookkeeping and official records, public reporting of budgets and political funding, official powers of supervision and intervention, or reports filed by citizens and whistleblowers – may, in the wrong hands, be used to mislead investigators or even to conceal the problem. Investigations and trials often become weapons in political struggles; where that is the case, apparent trends in corruption may say more about contention among key elites and factions than about the actual abuse of power or progress of reform.

### Notable efforts to measure corruption

52. A variety of **corruption risk measures** differing in coverage, methodology, and quality are now available. Some of the longest-running efforts at measurement have been mounted by firms providing risk assessments – at times, on a proprietary basis only – to international business. These have at various times included surveys by Political and Economic Risk Consultancy,<sup>6</sup> the Institute for Management Development,<sup>7</sup> Political Risk Services,<sup>8</sup> *The Economist* Intelligence Unit,<sup>9</sup> and Business International (now a part of *The Economist* group). Others are produced by advocacy groups such as the World Economic Forum<sup>10</sup> and Freedom House,<sup>11</sup> survey organisations including Gallup,<sup>12</sup> publications such as *The Wall Street Journal*, and expert panels, sometimes working in affiliation with international organisations. In the United States the Public Integrity Section of the Department of Justice regularly publishes data on corruption prosecutions and convictions in federal courts, reported by judicial districts.<sup>13</sup> Economists have

<sup>5</sup> This is what is generally meant by the somewhat awkward term “operationalisation”.

<sup>6</sup> [www.asiarisk.com](http://www.asiarisk.com); see “Corruption in Asia Report” at [www.asiarisk.com/subscribe/dataindx.html](http://www.asiarisk.com/subscribe/dataindx.html) (Viewed 5 September 2008).

<sup>7</sup> [www.imd.ch](http://www.imd.ch) (Viewed 5 September 2008).

<sup>8</sup> [www.prsgroup.com](http://www.prsgroup.com); data archives at [www.prsgroup.com/CountryData.aspx](http://www.prsgroup.com/CountryData.aspx) (Viewed 5 September 2008).

<sup>9</sup> [www.eiu.com](http://www.eiu.com) (Viewed 5 September 2008).

<sup>10</sup> [www.analystsforum.org](http://www.analystsforum.org) (Viewed 5 September 2008).

<sup>11</sup> [www.freedomhouse.org](http://www.freedomhouse.org) (Viewed 5 September 2008).

<sup>12</sup> [www.gallup-international.com](http://www.gallup-international.com) (Viewed 5 September 2008).

<sup>13</sup> The most recent report, for 2006, is available at [www.usdoj.gov/criminal/pin/docs/arpt-2006.pdf](http://www.usdoj.gov/criminal/pin/docs/arpt-2006.pdf) (Viewed 5 September 2008). (START).

used economic data that, while not measuring corruption *per se*, tap into closely-related problems such as "black-market premiums" (Bahmani-Oskooee and Goswami, 2004) or the quality of countries' institutions (the study that gave rise to a large literature on institutional quality was Knack and Keefer, 1995). A 1999 report by the United Nations Crime Prevention and Criminal Justice Division<sup>14</sup> compiling criminal justice data included statistics on bribery. The data encompass many countries and a long time span; on the negative side there are clear questions of comparability across diverse court and investigatory systems. More recently the United National Office on Drugs and Crime (UNODC) launched a series of national-level Crime and Corruption Business Surveys (CCBS)<sup>15</sup>. Trial projects have been carried out in Bulgaria and Canada in 2005 and 2006, and the first full national survey was conducted in Cape Verde in 2006. Useful counterparts to the UNODC work include the World Bank's "Doing Business" surveys<sup>16</sup> and the OECD's "Cutting Red Tape" project<sup>17</sup>. Both of those assessments focus upon administrative barriers and burdens impeding efficient and orderly market activities; such burdens, where excessive, are blinking indicators both of past corruption and of incentives sustaining it, as discussed below.

53. **The Opacity Index**<sup>18</sup> (Hall and Yago, 2000; Kurtzman, Yago, Phumiwaswana, 2004: 12; Kurtzman and Yago, 2008), by contrast, gets at corruption by way of its correlates and consequences, many of them measurable with some precision. It measures "the costs and preponderance of high-frequency, low-impact risks resulting from corruption, a lack of due process, poor regulatory enforcement, and non-standard reporting practices, each of which adds substantial costs to global business" (Kurtzman and Yago, 2008: 1). A statistical model incorporates data of corruption (Transparency International's Corruption Perceptions Index and the PRS Group's International Country Risk Guide<sup>19</sup>), efficacy of the legal system, "deleterious economic policy", inadequate accounting and governance, and detrimental regulation – and estimates the net effect of such factors upon interest rates (the "premium") paid by businesses operating in each of 48 countries. (The 2001 Opacity Index, by contrast, estimated interest premiums with respect to governments' sovereign debts). Not surprisingly, high estimated levels of corruption are associated both with unfavourable Opacity scores (interest rates are higher where corruption and related difficulties are more severe) and with difficulties of macro-economic development. The index remains a research enterprise rather than a hard-and-fast public ranking; its constituent data are far from perfect, and coverage is limited to a relatively small number of countries. But the Opacity Index is worth

<sup>14</sup> [www.uncjin.org/Special/GlobalReport.html](http://www.uncjin.org/Special/GlobalReport.html) (Viewed 7 November 2006) and the agency's *Global Report on Crime and Justice* (Oxford: Oxford University Press, 1999).

<sup>15</sup> United Nations Office on Drugs and Crime, "Crime and Corruption Business Surveys". Background information available at [www.unodc.org/unodc/en/data-and-analysis/Crime-and-Corruption-Business-Surveys.html](http://www.unodc.org/unodc/en/data-and-analysis/Crime-and-Corruption-Business-Surveys.html) (Viewed 8 September 2008).

<sup>16</sup> World Bank, "Doing Business" annual series of data and reports; background information available at [www.doingbusiness.org/](http://www.doingbusiness.org/) (Viewed 8 September 2008); data and reports available at [www.doingbusiness.org/Downloads/](http://www.doingbusiness.org/Downloads/) (Viewed 8 September 2008).

<sup>17</sup> Organization for Economic Co-operation and Development (OECD), "Cutting Red Tape: Comparing Administrative Burdens Across Countries". See [www.oecd.org/document/26/0,3343,en\\_2649\\_33735\\_39385178\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/26/0,3343,en_2649_33735_39385178_1_1_1_1,00.html) (Viewed 8 September 2008).

<sup>18</sup> The most recent report, for 2007-2008, is Kurtzman, Joel, and Glenn Yago. 2008. "Opacity Index 2007-2008: Measuring Global Business Risks." Santa Monica, CA: The Milken Institute. Available at [www.kurtzmandgroup.com/reports/opacity\\_index\\_2.pdf](http://www.kurtzmandgroup.com/reports/opacity_index_2.pdf) (Viewed 4 September 2008).

<sup>19</sup> See at [www.prsgroup.com/](http://www.prsgroup.com/) (Viewed 11 September 2008).

careful consideration because it is linked to real processes – the borrowing and lending of capital – that theory tells us should be quite sensitive to corruption and related issues, and that are pragmatically re-evaluated on a minute-by-minute basis. Those evaluations are made by market specialists with a vested interest in getting things right – in effect, by a very large if informal “expert panel” – and who receive continuing feedback as to how accurate their assessments have been.

### Perceptions and realities

54. Perception-based indices draw upon a number of sources. Some sample the public at large; others depend upon expert assessments. People involved in international trade and investment are among those most often included in such surveys; given the lack of harder data, the fact that much corruption arises in the context of business deals, and the extent to which those people move about the global economy, the approach makes some sense. Moreover, perceptions of corruptness are significant in their own right as they influence foreign policy, aid, investment, and lending decisions.

55. But those judgments reflect many influences besides direct knowledge of corruption: the overtness of corrupt dealings, personal reactions to societies, people and their behaviour, or investment stakes and the degree of success a firm has experienced in a given country might all skew perceptions. For example, mature market societies – where high-level relationships between state and markets are legally institutionalised and relatively predictable – might well receive a kind of “pass”, ignoring lower-level corruption affecting citizens and small businesses. Not all countries are included in all surveys; for some the underlying database is rich and varied, while for others only a small number of soundings are available.

56. Sample sizes vary widely, as do the questions and the comparisons they request. Some surveys have respondents rate overall levels of corruption on a scale while others ask about bribes, extortion, or other irregularities in specific governmental functions, or tap respondents' own experiences. Still others invite respondents to estimate the extent to which corruption or bribery harm a given country's economy, an issue far too complex to be a matter of opinion. Such questions also invite the conclusion that because country X is poor it must also be highly corrupt, or, that wealth by itself indicates good governance. Contrasting types of corruption, the stakes involved, international sources and connections, effects upon specific segments of the population or economy, and connections to organised crime or violence are usually omitted in such surveys. Results invite comparisons, but it is difficult to say what is being compared, or what sorts of variations are exaggerated or “flattened out” across a range of scores. Comparisons over time are even more suspect, an issue to be examined in more detail below.

57. The best-known international corruption index, **Transparency International's Corruption Perceptions Index** (CPI),<sup>20</sup> exemplifies this approach. Launched in 1995 and updated annually, the CPI is a “poll of polls” averaging the results of a range of public opinion surveys, expert rankings (the 2007 edition was based entirely on expert panels), and other sources into a zero-through-ten ranking of countries, with a high score indicating a low level of perceived corruption. It has won worldwide attention, sparked myriad analytical studies, and – perhaps most important – has helped sustain public and official attention to corruption as a critical issue of development and governance. Coverage has expanded from 41 countries in 1995 to 179 in 2007, and methodology and documentation have been refined on a continuing basis. Rankings are based on averaged results from a minimum of three surveys, and from as many as ten in the case of India and eleven for Indonesia. Over time the CPI has spawned related efforts by TI such as a “Bribe Payers' Index”,<sup>21</sup> which purports to compare the “supply-side” willingness of private parties in

<sup>20</sup> Data, documentation, and press kit available at [www.transparency.org/policy\\_research/surveys\\_indices/cpi](http://www.transparency.org/policy_research/surveys_indices/cpi) (Viewed 8 September 2008).

<sup>21</sup> See [www.transparency.org/policy\\_research/surveys\\_indices/bpi](http://www.transparency.org/policy_research/surveys_indices/bpi) (Viewed 8 September 2008).

various societies to pay bribes elsewhere, and the *Global Corruption Barometer*,<sup>22</sup> offering more detailed evidence of public perceptions of corruption in key institutions of various nations (the CPI and its methods are discussed by its creator in Lambsdorff, 2006, and Lambsdorff, 2007; for a critique of perception-based measures employing data from the *Global Corruption Barometer* itself, see Abramo, 2007; for further analysis of the differences between perception and experience, see Olken, 2006, and Gatti, Paternostro, and Rigolini, 2003; and for an argument that the CPI in its present form has outlived its usefulness, see Galtung, 2006).

58. An even more ambitious measurement effort is the World Bank Institute's **Worldwide Governance Indicators** (WGI), part of the larger "Governance Matters" project.<sup>23</sup> A wide variety of data, including most of the survey data incorporated into the CPI but others as well, are used to estimate six attributes annually for over 200 countries and territories, namely:

- Control of corruption;
- Voice and accountability;
- Political stability;
- Government effectiveness;
- Regulatory quality; and
- Rule of law.

59. The dataset also includes useful links to over thirty sources of constituent data, although some of the underlying data are proprietary and at least one source is confidential. The WGI data are calculated using a sophisticated "unobserved components" approach (Greene, 1990). They are notable for their extensive coverage (over 200 countries and territories) and for breaking governance out into six components that embody diverse positive values. Questions have been raised, however, as to what if anything the data actually measure, and about the ways errors are dealt with in the modelling and calculation processes (Thomas, 2007). Arndt and Oman likewise point out a variety of problems with the WGI data, identifying major difficulties in terms of error; comparisons over time; sample biases and a resultant "echo chamber" problem reminiscent of error issues identified by Thomas; and transparency of data-gathering and computation (Arndt and Oman, 2006: Ch 4; for a defence see Kaufmann, Kraay, and Mastruzzi, 2006, and for a discussion of next steps in the refinement of such indicators, see, Kaufmann and Kraay, 2008).

### Methodological problems

60. Some of the problems of corruption indices are methodological, and are grounded in basic issues of measurement. In principle analysts can measure anything (Babbie, 1995: 110). But that is more easily

<sup>22</sup> See [www.transparency.org/policy\\_research/surveys\\_indices/gcb](http://www.transparency.org/policy_research/surveys_indices/gcb) (Viewed 8 September 2008).

<sup>23</sup> D. Kaufmann, A. Kraay, and M. Mastruzzi, 2008. "Governance Matters VII: Governance Data for 1996–2007", World Bank Institute paper; available, along with data and a variety of other sorts of documentation, at Kaufmann, D., A. Kraay, and M. Mastruzzi. 2008. "Governance Matters VII: Governance Indicators for 1996–2007", World Bank Institute paper; available, along with data and a variety of other sorts of documentation, at <http://info.worldbank.org/governance/wgi/resources.htm> (Viewed 5 September 2008). (Viewed 5 September 2008).

said than done, and much can be lost (and a great deal of noise can be added) *en route* from essential concepts and nominal definitions to the events or artifacts included in operational measures. Many concepts are categorisations of, or inferences from, phenomena that are difficult to specify in advance, such as “democracy” (Collier and Levitsky, 1997). We know it when we see it, but the concept remains essentially contested (Gallie, 1965). Over time the concept “creeps” away from its starting point (Collier and Levitsky, 1997) or is re-defined *de facto* in terms of consequences or common practices such as elections. The original idea, as well as the full range of variations and complexity in its actual manifestations, may well be lost. A more subtle danger is **reification** (Babbie, 1995: 116-118) – thinking about operational measures as though they were the concept itself. Too often analysts end up studying things because they are easily counted, or treating a flawed indicator (such as a corruption perception index) as a measurement of a corruption itself, simply because the figures are available.

61. Three criteria – highlighted in the following box – by which analysts and decision makers judge any measurement are:

- **Validity:** do data measure what it is claimed they do?
- **Reliability:** do repeated applications of a measurement return consistent values?
- **Precision:** by what units or categories – fine or coarse – should results be expressed?

**Box 1. Measurement criteria: Validity, reliability and precision**

**Validity:** do collected data actually measure what analysts claim they do? Some “measurements” draw on something in addition to, or other than, that which they claim to measure – or upon nothing at all. Concepts themselves do not exist in the real world, or have “real definitions” (Babbie, 1995: 116; Hempel, 1952). They are only constructs useful for categorising objects or events, drawing out attributes analysts think they share. Empirical measures can never be better than approximations, and useful simplifications, of complex realities, and even the most careful definition is “nominal” only – expressed in words that are only approximations of those approximations.

Analysts can assess validity in several ways (Babbie (1995: 127-128). Does a measure possess **face validity** – that is, does it have anything to do with the concept? Plugging climate data into a model of corruption might produce significant correlations, but analysts would not be measuring “corruption” or any other reasonable factor. Does it possess **criterion-related** or **predictive** validity, in the sense of predicting changes in other variables that theory tells us should be related? Or, a measure might be related to others in ways that are consistent with what analysts know about them, even if it does not “predict” them – an attribute called **construct validity**. Corruption measures, for example, should return high values where institutions are weak (Knack and Keefer, 1995). A measure with **content validity** works well across diverse aspects of a concept: corruption data should reflect all major varieties, not just bribery. Finally, a concept might have **reference-group validity** – that is, be judged sound by people with extensive knowledge of what analysts wish to measure.

**Reliability:** does measure return consistent results? A corruption scale giving a country an eight out of ten one year, two the next, and five the year after that likely has low reliability. Equally useless would be a rating system producing divergent scores when applied to the same case by different observers. No measure will be completely reliable; indeed, what looks like strong reliability might in fact be a validity problem, manifested as insensitivity to change.

**Precision:** how finely should a measure be divided and expressed? Usually more precision is better: a “yes/no” corruption variable would be useless. High-, medium-, and low-corruption categories would be better, and numerical rankings more precise yet. But there can be false precision: it is more useful to know that a country’s population density is 255 people per square mile than to say that it is “moderate”, but neither useful nor statistically appropriate to express that measure as 255.34890634. One measurement can be more precise but less accurate than another: data claiming Country X’s population density is 255 people/mi<sup>2</sup> may be less accurate than a ranking of “moderate” if the true figure is 75 people/mi<sup>2</sup>.

Source: Babbie, Earl. 1995. *The Practice of Social Research*.

62. Clearly, validity is a core problem with existing corruption indices. Disagreements over definitions of the concept are just one problem. More to the point, perceptions are not the same thing as corruption itself: as noted they may reflect a wide range of influences, or types of “noise”, unrelated to the extent and seriousness of corruption in a society. Perceptions do shape important decisions, but the danger is that they will lead to an “echo chamber” problem (a detailed discussion appears in Thomas, 2007) in which officials and investors repeat what they hear from each other, giving anecdotes false authority through repetition. Corruption indices exhibit impressive reliability in some respects, with year-on-year results often correlating at upwards of +.95, and such consistency is sometimes advanced as evidence that the indices measure something real. But reliability is not evidence of validity; indeed, results at such a high level of consistency, year after year, may be artifacts of underlying methodology (such as the re-use of some survey results), show an insensitivity to actual trends, or even be evidence of the echo-chamber phenomenon itself. As for precision, it is possible to have too much of a good thing. One critique of the CPI in its early years was that by reporting results out to two decimal places – scores were given as, say, 3.87 and 7.22 – scores implied a precision greater than the number of significant digits justified by the underlying data. It might be more accurate to group countries into broad bands – high, high-medium, medium, etc. – even though such results would be less intuitively satisfying not so easily plugged into statistical models (see the discussion of levels of measurement in the box that follows).

#### **Box 2. Levels of measurement**

Measurements can differ in terms of **levels of measurement**. Higher levels contain more information and can be used in more elaborate statistical treatments, but analysts often see information at one level of measurement discussed or analyzed as though it were at another, usually higher, level.

Some measures are **nominal**, grouping cases into named categories that each embody some common characteristic, but among which no overall relationship exists (an example would be a list of public employees grouped by region or municipality). Others are **ordinal**, grouping cases into categories that again reflect a shared characteristic, but that can also be ranked higher or lower in terms of some attribute common to all. Analysts might, for example, place countries into high, middle, and low-income groups; all in the “high” category would be more affluent than those in the “middle” group, but there would be considerable variation *within* groups and no assurance that the differences *among* groups are the same.

**Interval**-level measurements array cases along a common dimension demarcated in units of identical size, but without a point indicating the complete absence of the attribute being measured. The Fahrenheit scale is an example: all values refer to a common attribute (heat), and a one-degree difference is identical across all values, but as its zero point is arbitrary a reading of 60 degrees cannot be interpreted as twice as warm as 30. An agency might survey citizens or clients, asking them to rate their satisfaction with its services on a minus-five to plus-five scale; while some ratings might be expressed as zero, that would not refer to a complete absence of satisfaction (and, as with the temperature example, a rating of 4 would not suggest twice as much satisfaction as a score of 2).

Finally, **ratio**-level data also array cases along a dimension marked off in units of identical size, but possess a true “zero point”; thus, expressions of proportion are appropriate. A country with 50 million residents is twice as populous as its neighbour with 25 million. Level of measurement is an important issue: it is tempting to treat ordinal data as interval-level, for example, but the results can be misleading.

63. Corruption indices, and the scholarship and debates they have spawned, generally support much of what analysts had long suspected. Scores are worst for poor, undemocratic, and unstable countries. Multivariate analysis indicates that corruption is both cause and effect of a variety of economic and

political pathologies.<sup>24</sup> But when analysts move from an overall estimate of corruption across entire countries to actual reform and integrity-building for the long term, the requirements for feedback become considerably more demanding.

### Practical difficulties

64. How much guidance do corruption indices give reformers? By way of illustration, consider what may well happen to single-number perception scores for a country that has begun to take meaningful measures against corruption? At the very least, progress will be uneven, as will society's awareness that progress is occurring at all. In that setting, the uncertainty of scores may widen considerably, while changes in whole-country scores will be difficult to interpret. An effective anti-corruption campaign will produce revelations of wrongdoing, convictions, and countless new allegations. This is all the more likely in a democratising country with citizens, journalists, and opposition figures feeling more free to speak out, and contending factions using corruption allegations to settle old scores. Citizens, officials, and anyone asked about perceptions of corruption may well have trouble distinguishing the bad news from the good. Reforms that break up corrupt networks may well lead to a surge of overt, smash-and-grab corruption as desperate elites take as much as they can, as fast as they can take it (Scott, 1972; Knack and Keefer, 1995). Perception scores could well *worsen* markedly, at least in the short run – as a consequence of effective reform.

65. A related problem is that perception-based corruption measures could ironically become self-fulfilling prophecies (Heywood, 2008; Williamson, 2004): that is, a reputation for extensive corruption could drive out much-needed investment and valuable international partnerships – the latter, in both public and private sectors. The resulting losses in aid, trade, infrastructure, capital, and engagement with the broader global economy (on this last point see Treisman, 2000) will only make reform more difficult. Such a reputation could also lead to more intense interest in a country's scandals and corruption allegations, depending upon levels of economic or strategic interest in the country, on the part of both international and domestic news media, resulting in more reports of corruption and, possibly, even worse perception scores. Such intensified attention would be wholly appropriate **if analysts are confident that unfavourable scores actually signal worse corruption problems**; but there are reasons to be cautious about such indices. Indeed, a country singled out as highly corrupt might become more attractive to the wrong sorts of entrepreneurs, or to criminal groups, searching for a safe haven for unsavoury activities. For these reasons analysts must always remember that whether or not perceptions are useful in measuring corruption, gathering and publishing them in league-table form is an exercise in labelling as well – one to be undertaken only with great caution.

66. Current indices leave us with three serious practical problems:

- **Most existing indices cannot track trends over time.** CPI results are affected by the number and identities of societies being compared, as well as by annual changes in available data and the ways they are gathered. Those data are assembled, sometimes at long and irregular intervals, by different bodies with contrasting methods, budgets and agendas, and in some instances are re-used in the calculation of index scores for several years running. Societies differ in terms both of public knowledge of corruption and the presence of international organisations and businesses, and thus in the extent, speed and accuracy with which actual changes might become public knowledge. In any event it is unlikely in any society that *all* trends in corruption actually run in the same direction at the same pace; even if analysts had accurate country-wide data on changes over time they would be of relatively little use to those in charge of reform.

---

<sup>24</sup> A large bibliography emphasizing studies of the developmental implications of corruption can be found at <http://people.colgate.edu/mjohnston/Total%20Bib%20Fall%2005.doc> (Viewed 8 September 2008).

67. Analysts know little about “lag times” between real changes in corruption and trends in perceptions: indeed, in some societies such connections may never be made, or may be overridden by rumours and officially-contrived distractions. Different sorts of events and changes seem likely to affect perceptions in different ways: major high-level scandal might move public opinion more quickly, and in more lasting ways, than more pervasive if less dramatic lower-level abuses. The gradual progress of specific local or agency reforms might be even less likely to affect index scores. Events are open to differing perceptions: the conviction or resignation of a major figure will be seen by some as signifying a major surge of corruption, and by others as evidence that reform is finally taking hold. Because of more elaborate scoring and computational methods, data that are not mapped onto a closed-ended scale (such as the CPI’s 0-10 rankings), and country coverage that is more consistent year after year, results in Worldwide Governance Indicators may be somewhat better than the Corruption Perception Indicators at tracking changes over time, but Arndt and Oman (2006, Ch. 4) list strong reasons to be cautious with those data as well.

- The “**single-number problem**”: country-level indices compress differing types of corruption occurring in different segments of societies and governments into one-number scores (for a detailed critique of single-number assessments, see Arndt and Oman, Ch. 4 *et passim*). Comparisons along the one-dimensional indices that result assume, in effect, that corruption is the same thing wherever it occurs, is distributed uniformly across all regions and institutions of a society, and varies among countries only by degree. Statistical explanations of index scores apply the same substantive explanations everywhere. But neither corruption nor reform is a single national-level process. Each is a series of complex, highly specific, and often interactive steps to alter the ways wealth and power are pursued, used, and exchanged. Corruption is diverse and changeable; there are many forms (Johnston, 2005c) and contrasts within most societies. How much nepotism or patronage is equivalent to a certain level of bribery in road construction? Is that bribery comparable in importance to similar practices in arms contracting? Is corruption that is closely linked to violence simply a more serious case of the same thing that occurs when parking tickets are “fixed” or campaign contributions are misused? No national score accurately reflects variations between Northern and Southern Italy, across Russia, or among Minnesota, Alabama, and New Jersey. Some countries have high-level corruption, others find it lower down the political or bureaucratic hierarchies, and still others see most abuses in electoral politics and patronage. Sometimes the problem centres around international trade or transnational crime, while in others cases it is home-grown.

68. Even where corruption has become endemic or systemic it feeds upon specific government functions, programs, and agencies. Reform requires “micro” knowledge of the incentives and constraints affecting behaviour, and accurate feedback at those same levels of detail. Further, it is essential to demonstrate the effectiveness of reforms to citizens, political leaders, public employees, and would-be participants in future corruption. No single number can serve all of those purposes.

- **From assessment to action:** a third problem – one implied by a useful distinction between outputs and outcomes – is more subtle but equally troubling. Integrity does not reduce just to an absence, or significant decline, of corruption. Corruption, even where it is severe, does not explain all that goes wrong in a government or society, nor does it negate all that is good. Policies may be poorly conceived and under-funded, the current state of knowledge may be insufficient for the problems policy makers and managers are tackling, and even the best governments are rarely nimble enough to keep up with the pace of social change.

69. Knowing that corruption is extensive does not, by itself, tell policy makers and agency managers what to change and what to leave alone. Integrity assessments must:

- Tap into the positive values analysts and officials seek – justice, a high quality of life, responsiveness and accountability in government, the efficient use of resources, and delivery of quality services.
- Be actionable, an awkward term for an important idea: the best assessment standards will point to specific vulnerabilities and countermeasures, rather than just reminding us that governance is generally out of joint.
- Finally, accurately reflect progress, or lack of it, if analysts, policy makers and managers are to know with any confidence whether reform efforts are succeeding.

70. Thus, existing indices are likely to help us least in the countries with the worst corruption problems, *and* in those that are most aggressively tackling their corruption problems – the two groups of societies that are among the most pressing concerns. Perceptions may outrun, or lag behind, actual trends. Any comprehensive anti-corruption strategy will likely work better with some varieties of the problem than with others, and yet a single-number index will not be able to tell us much about those contrasts, where to attack first, or which aspects of a strategy are working and which are not. These are much more than mere methodological niceties. Analysts – and even more so, managers and reformers – need valid, reliable, and tolerably precise data of problems and progress that draw, as much as possible, upon objective data (a useful discussion of options appears in Arndt and Oman, 2006: Ch. 6).

#### IV. BUILDING AND SUSTAINING AN INTEGRITY FRAMEWORK

71. How, if analysts do not seek to construct yet another corruption index, can vulnerabilities to such a diverse range of threats be assessed, and progress toward serving such an ambitious values be tracked?

72. Many previous and current efforts conceive of the problem primarily in terms of corruption control, and see the problem essentially as one of deviance requiring various sorts of crime-prevention approaches to reform. Another approach is to focus on perverse and counterproductive incentives – like corruption control, clearly a legitimate concern, but one that tends to emphasise specific tasks and outputs rather than the full scope of integrity. Such approaches are not so much wrong as incomplete: they tend to isolate government from societal influences, emphasise individual failings or offenses, and put relatively little emphasis upon the outcomes of government.

73. The OECD’s Ethics Infrastructure and the updated Integrity Framework (described in detail in OECD, 1996, 2000, 2005, and updated in OECD, 2008), by contrast, brings a renewed emphasis to two major ideas. The first is the encouragement of **positive goals and accomplishments**. The second is the idea that just as threats to integrity can become institutionalised in, and even sustained by, official structures and processes, **integrity must be institutionalised and sustained too**. The first notion is important if policy makers and managers are to do more than just react to threats with new rules and punishments, and have anything other than lists of attempted reforms by which to judge progress. The second is essential because integrity is only partly related to “being good”. It has much more to do with establishing a system and culture of integrity within organisations, government operations, and society itself that expects and rewards justice, accountability, efficiency, and transparency. The sorts of self-interest that might otherwise drive corruption can then help sustain reform.

74. An effective and sustainable Integrity Framework requires a feedback loop, however. How can analysts, policy makers and managers assess integrity, identify vulnerabilities, and demonstrate progress in credible, widely understood ways? Can analysts, policy makers and managers properly judge which elements of the framework are working effectively, which are ineffective, and which might even be counterproductive? As a later section will show, attempts at such assessments by measuring governance and corruption are numerous but, in some cases, of doubtful value; most are far too general, and remote from actual integrity problems, to be useful. Rather than using these numbers because they are there, it is appropriate to start anew, and to recognise that useful assessments must be:

- **Detailed:** whole-country index numbers – even if they are credible in themselves, which some are not – are too general.
- **Objective:** assessments must be based on verifiable evidence rather than on opinions, and expressed in actual units rather than points on arbitrary scales.
- **Non-invasive,** with respect to agency operations – that is, the act of measurement must not change or bias that which is being measured – and with respect to officials’ personal rights.
- **“Policy-neutral”:** measurements should not encourage some substantive policies or discourage others (Manning, Kraan, and Malinska, OECD 2006). As a practical matter, the co-operation of governments, political leaders and officials, essential to the DB strategy to be proposed here, will be undermined or pre-empted if policy neutrality is not maintained.

- **Low cost:** particularly because repeated assessments over time are essential, measurement strategies must be inexpensive, both in monetary terms and in terms of keeping additional data-gathering and -reporting requirements to a minimum, and also cost-effective to yield added value. High-cost methodologies are likely to be seen as invasive and unlikely to be repeated as a matter of routine.
- **High in validity and reliability:** any technique must measure what analysts claim it measures (validity), and return consistent values across multiple applications; these issues will be discussed in more detail below.
- **Transparent and easily understood:** not only analysts and officials, but citizens and civil society groups must be able to interpret integrity assessments easily and accurately. The best measurements will be interpretable in terms of positive values.
- **Trackable over time:** for reasons to be discussed, existing corruption measurements offer little or no valid information about change over time. Analysts and officials must be able to demonstrate progress, or lack of it, to citizens, political and administrative supporters, and potential wrongdoers, and successful leaders and managers must be able to claim credit for their accomplishments.
- **Supportive of trust:** integrity measures cannot succeed if they foster distrust within government or between it and citizens, or if they create an environment in which officials feel someone is constantly looking over their shoulders. Measures that directly assess levels of mutual trust and the forces supporting or threatening it will be essential.
- **Actionable:** assessments should not only tell us that a situation is bad or good, improving or deteriorating, but should also point directly to improvements and reforms that are likely to succeed – steps that are “actionable” at a strategic level, rather than focused just on individual officials and cases.

75. Assessments meeting these criteria can offer a detailed picture of trends in integrity. Clear connections between cause and effect will help us answer to the most persistent question faced by any effort to improve governance: that is, “what works?” Further, as Heintzman (2005) rightly points out, an emphasis on **outcomes**, and not just **outputs**, is essential in the value chain. By that he directs the attention not just to specific things governments do, but also to the broader wellbeing of society and its citizens. Policy makers and managers need to know whether citizens, and justice itself, are well served, policies are faithfully implemented, and public power and resources are used in honest and accountable ways. To move agencies and officials toward those sorts of outcomes is not only to improve the quality of government; it is a good thing in itself.

76. A final consideration is co-ordination to ensure consistency with ongoing OECD efforts – in particular, supporting the construction of sound *Integrity Framework* and the assembly of the first and subsequent editions of *Government at a Glance*. The former provides, among many other things, a useful map of integrity efforts, identifying three major domains within which data and benchmarks must be gathered and among which knowledge can be shared. The latter deals more specifically with data and the ways they fit into a broader understanding of government.

### **The Integrity Framework: Domains of assessment**

77. Defining integrity and its core values is one challenge – one that has scarcely been laid to rest here. But where, in practice, do analysts, policy makers and managers search for evidence?

78. Many efforts over the past decade have focused upon attempts to measure corruption – clearly one element of integrity assessment, but only one – assessed by surveys of perceptions that are tabulated at the national level. Others have broadened the focus somewhat: for example, the World Bank Institute’s Worldwide Governance Indicators consider six dimensions, including not only corruption control but also regulatory quality, voice and accountability, and three more. But most of the underlying data there, too, rest on perceptions, and results are presented as scores for whole countries. For reasons to be discussed below, both the perceptions approach and assigning single scores to whole countries are problematical practices.

79. Another drawback is of greater concern at this point: corruption assessments, even if they draw upon a broad database within a country, are aimed at one kind of problem, or one set of vulnerabilities, at the level of governmental processes. Important as those concerns may be, integrity involves much more. The **integrity climate** – values, morale, clarity of ends and means, to name a few aspects – shapes processes “from within”. Stated procedures may mandate certain kinds of conduct while forbidding others, but if employees are un- or under-paid, morale is low, lines of accountability are unclear (or used to perpetuate misconduct, as is sometimes the case), and the agency’s goals are vague, integrity will suffer. Likewise, relationships between government and society, including outcomes of government activities, openness and responsiveness of institutions, and public responses will have much to do with whether, and how effectively, integrity is demanded and rewarded.

80. Experience confirms how sensitive integrity management is to implementation deficits: when ambitious plans are not implemented or instruments and procedures are not applied in daily work. Implementation of measures for fostering integrity and preventing corruption take place on organisations level. Consequently, a sound Integrity Framework (OECD: 2008)<sup>25</sup>, principally aims to focus on the instruments, processes and structures that foster integrity and prevent corruption and other integrity violations within public sector organisations. The components of the framework support the following four functions:

- Determining and defining integrity;
- Guiding towards integrity;
- Monitoring compliance; and
- Enforcing integrity.

81. The Integrity Framework takes a systemic approach to ensure that its components are mutually enforcing each other as they are interdependent. Although core integrity instruments (e.g. code of conduct, conflict-of-interest policy, whistle-blowing procedures, etc.), core processes and actors (e.g. integrity officer or special ethics unit as “integrity actors”) with primary aim to enhance integrity are vital, however, these are not sufficient themselves to achieve the goals of strengthening integrity in public sector organisations. The Integrity Framework as a whole is also dependent on its wider management and governance context. Therefore, it involves processes and actors – such as personnel management, financial and contract management – that are typically part of adjacent management fields and they do not focus primarily on integrity (e.g. the primary aim of procurement and contract management is value for money). Core and complementary integrity instruments, processes and actors together with the context – instruments, process, actors within and outside public sector organisations – that influence the integrity of

<sup>25</sup> *Towards a sound Integrity Framework: Instruments, mechanisms, actors and conditions for implementation*, OECD (GOV/PGC/ (2008)21) September 2008; *Elements for an updated framework for integrity and corruption prevention in the public service*, OECD (GOV/PGC/ETH(2007)1), June, 2007.

members of an organisation constitute the Integrity Framework. The following table highlight the structure and relations of key components of the Integrity Framework.

**Table 1. The Integrity Framework**

Integrity Framework							
Integrity Management Framework					Context		
Core measures			Complementary measures			Inner context	Outer context
Instruments	Processes	Structures	Instruments	Processes	Structures		

82. The DB strategy proposed here gathers evidence in the following key domains of the Integrity Framework defined above:

- The **core integrity instruments**, processes and actors can be assessed in terms of existence, implementation and functioning of values and standards, guidance to support implementation and controls to monitor and ensure compliance with values and standards of expected conduct. Laws, codes, institutions, procedures for guidance and monitoring, and incentives are part of the mix, and their presence or absence is relatively easily verified. Less clear but equally important, however, is the internal “climate of integrity”: factors such as commitment to established standards; follow-through on the use of existing standards or procedures for control; the extent to which misconduct and other breaches of integrity continue to occur, and perhaps to be tolerated or expected; whether or not the knowledge acquired through the application of integrity measures is analyzed and applied to further efforts; the extent to which ends, means, and lines of accountability are clearly understood, and so forth.
- **Analysing risks to integrity** is the cornerstone of the Integrity Framework. Verifying resistance to corruption is best approached via collecting evidence on vulnerabilities that, is “officials in situations and/or positions at risk”. Those vulnerabilities, in turn, can be measured through carefully-chosen and – understood data on the processes and outputs that shape incentives to corruption. As will be explained, there are numerous sorts of data on what officials and agencies do, and how they do it, that if “benchmarked” against appropriate norms can tell policy makers and managers about the cumulative effects of threats to integrity and the incentives that check, or exacerbate, those threats and vulnerabilities.
- The context of the Integrity Framework **supportive public management and governance conditions** incorporates a wide range of factors. This also addresses ways citizens, officials, and agencies understand and respond to the outcomes of government activities, and the extent of constructive interaction that takes place. As suggested above in the discussion of integrity itself, outcomes that consistently reflect, and advance, the values and preferences of societies require, and are reliable indicators of, high levels of integrity in all of its several respects. They are also, of course, good things in themselves, and the basic reason why integrity assessments are worth doing in the first place.

### Providing evidence: Government at a Glance

83. Quite sensibly the *Government at a Glance* initiative will not make reference to government “performance”. That notion, while tempting and commonly used, is an oversimplification that may conceal more than it reveals. Even if they refer to specifically quantifiable results or rely upon surveys of public satisfaction, without specifying antecedent components “performance” data tell us little about why things are or are not going well (these issues are discussed in much greater depth in Manning, Kraan, and Malinska, OECD 2006). Instead, *Government at a Glance* intends to emphasise five kinds of data (Manning, Kraan, and Malinska, 2006: 7-9 and *passim*):

- **Inputs:** resources, materials, investments, and other factors used to produce a service.
- **Public sector processes:** “structures, procedures, and management arrangements”.
- **Outputs:** results of services and processes, measurable in terms of quantity and quality).
- **Outcomes:** impact and consequences of processes and outputs for the community.
- **Antecedents and constraints:** the broader context of resources, constraints, and historical conditions within which all of the above take place and are assessed

84. The DB strategy proposes to primarily rely upon the first four kinds of data, and will incorporate the fifth by implication as outcomes at one given time shape, and become, antecedents and constraints affecting subsequent rounds of assessment.

### Sequencing: Implementing the strategy in phases

85. Ultimately all four main varieties of evidence – data on inputs, processes, outputs, and outcomes – will be involved in a detailed assessments identified above in the Integrity Framework that addresses resistance to corruption and supportive public governance and management conditions. Getting to that point, however, must be a phased process, beginning with the area in which the most evidence and research are likely to be in hand on the existence and functioning of instruments, in order to minimise the costs and disruption associated with launching the DB strategy, and to make the incentives and rewards of participating clear from the outset.

86. The **sequence proposed**, and the broad data-gathering and benchmarking activities implied are as follows; more detail on each phase follows in later sections of this paper:

- **First**, examining core components of the *Integrity Framework* via updating inventories of **instruments, processes** and **actors** for determining and defining integrity, guiding towards integrity, monitoring and enforcing integrity.
- **Second**, examining selected **inputs** and **public sector processes**, gathering data of a variety of management processes within public organisations (*e.g.* measures for fostering integrity in procurement and contract management, financial management, personnel selection and promotion, etc.), and benchmarking them against data from comparable jurisdictions. This phase of assessment is also aimed at measures for enhancing **resistance to corruption**, since (as will be explained in the following section) “blinking indicators” identified can tell policy makers a great deal about the past effects of threats to integrity, and about incentives that place officials and processes at risk After consolidating data on inputs and processes, data collection may review output data (*e.g.* statistics in procurement process).

- **Third**, gathering complementary data on less tangible, but quite important, aspects of inputs and processes. For example reviewing existing data related to “integrity climate” within public sector agencies, in order to connect them. Such surveys could provide procedural evidence on functioning and impact of core integrity instruments such as whistle-blowing policies and practices (e.g. whether procedures for “whistle-blowing” are accessible, protection is available, employees are committed to report misconduct, etc); integrity training and advice, etc.
- **Fourth**, repeated assessments in all three aspects to establish trends and underscore the importance and legitimacy of the DB strategy and the integrity issues it raises. Results of these assessments should be routinely reviewed first within public organisations but also published and opened up for public discussion; they can and should become the basis of consultation with citizens. Ultimately, these comprehensive assessments may allow policy makers and managers to build data-drive systemic models of the integrity situations within entire governments, to identify hitherto-unrecognised synergies or contradictions among the elements being assessed, and to recommend systemic improvements.

### “Blinking indicators” and responses

87. Underlying all of these phases is continuing attention to results, and specific steps to improve those results in functions and subgroups that emerge as problem areas. Properly constructed, many types and data and benchmarks should point directly to needed improvements and give a rough sense of how urgent a problem is. The most urgent of these can be thought of as “blinking indicators” – based on the consensus reached at the Expert Group – clear-cut signs of where integrity vulnerabilities are most pronounced.

88. The value of those sorts of indicators should not be underestimated. At present relatively few kinds of evidence on integrity give political leaders, agency managers, or citizens much of an indication of what should be done, or of which areas must be addressed first. At best, reformers find themselves reacting to past cases of corruption or other integrity failures, rather than heading off the most pressing current vulnerabilities; more often they confront governance situations perceived as defective in diffuse ways, but with little beyond good intentions and short-lived public outcry to guide their choices. In that sort of situation, demonstrating effectiveness and progress over time becomes very difficult indeed, and as a result public and political commitment to reform often fades away – only to become more difficult to mobilise next time around.

89. By contrast, where the data point to unreasonable delays or excessive costs in a particular program or process, that is evidence of vulnerability *and* an indication of actions to be taken. Improvements – faster processes, reduced costs – are signs of progress that can help sustain commitment to reform. The DB strategy provides relevant information to all concerned with pursuing positive values of integrity, not just broad indications of perceived failings, and allows those who address vulnerabilities successfully to claim credit for doing so. Failing responses to vulnerabilities can also be clearly identified, and those who have not succeeded in dealing those problems can be called to account – an outcome more likely to sustain pressures for integrity than diffuse survey data suggesting that government as a whole remains ineffective.

90. How these phases will work, and how analysts, policy makers and managers can draw such detailed inferences from data and benchmarks, is best made clear by a detailed discussion of each phase. The following chapter begins by examining the sorts of input, process and output data, and basic ideas about benchmarking, related the core components of the Integrity Framework.

## V. PHASE I: THE INTEGRITY FRAMEWORK

91. Providing data on the existence and functioning of key instruments, procedures and actors in place for determining and defining integrity, guiding towards integrity, and monitoring and enforcing compliance could support policy makers and managers build and sustain a sound integrity framework in public organisations. Collected data could focus, in particular on conflict of interest and post-employment measures, code of conduct or code of ethics, gifts and gratuities policy to provide benchmarks for implementation and impact. Decision makers and managers may also consider essential information on control and accountability processes, both internal (such as accessible channels and protection for whistleblowing, risk mapping) and external (*e.g.* providing efficient complaint procedures for citizens, scrutiny by civil society organisations, the press, and political leaders) affect specific actions by individual officials and clients and help define a climate of values, norms, and expectations for the longer term. So do officials' own experiences and judgments: how much wrongdoing or cavalier treatment of citizens an official sees on a regular basis; whether or not that the organisation's goals are clear, valued, and attainable; whether or not officials believe they and their work are valued by superiors and clients; whether the job brings with it meaningful status and a dependable living wage: these too are aspects of building an agency's *Integrity Framework*.

92. All of these factors are familiar to reformers, agency managers, policy makers and analysts of public administration. More can be done, however, to incorporate such variables into overall integrity assessments:

- A first step is to gather information on both sorts of influences identified above: formal integrity instruments and accountability procedures, and data on officials' own experiences.
- A second step is to relate them to each other through repeated regular measurements: how do agency procedures, their enforcement, and changes in them influence officials' views and experiences regarding their roles and work?
- Third – a benchmarking step – how do those results compare to those found in other agencies within the same jurisdiction, or in comparable bodies elsewhere?
- A final step is to incorporate data and benchmarks on the Integrity Framework into a comprehensive integrity assessment of all relevant domains outlined above.

93. Over time analysts, policy makers and managers will be able to relate trends in internal and external integrity measures, officials' own views and experiences, corruption vulnerabilities, and the broader social setting (to be discussed below) to integrity levels and problems as a key aspect of governance outcomes.

### Key elements of the Integrity Framework

94. As suggested above, data and benchmarks for the Integrity Framework focus on two critical kinds of elements.

95. First are the **formal elements of integrity management system**. These core instruments constitute the actual rules and tools of intervention that are used to shape integrity and resistance to corruption in the desired direction in public organisations. They support the following four functions, namely:

- Determining and defining integrity;
- Guiding towards integrity;
- Monitoring compliance; and
- Enforcing integrity.

96. These **core instruments** and institutionalised procedures and relationships aimed at ensuring high levels of integrity, and at detecting and reacting appropriately to gaps and infractions. Some of those instruments, procedures and relationships deal with behavioural aspects of integrity:

- Codes of conduct or codes of ethics;
- Conflict-of-interest rules, financial disclosure requirements;
- Gifts and gratuities policy;
- Post-employment measures;
- Reporting channels and protection for whistle-blowing;
- Internal integrity training, re-training, accessible advice and counselling;
- Internal integrity actors for co-ordination and external integrity “watchdogs” as supreme audit institutions and inspectors general;
- External evaluations of integrity; and
- Publication and public discussion of integrity issues, assessments.

97. Other **complementary management instruments** and processes focus more upon the ability of a public agency, or of subgroups within it, to perform mandated functions, however they also have an impact on integrity:

- Control mechanisms related to funds, materials, data, etc, in particular identified areas of risk such as public procurement, contract management and financial management;
- Role of integrity in personnel management, in particular in recruitment, evaluation and career promotion;
- Supervisory processes and relationships;
- Quality management, including measurement of activities and, where appropriate, output by officials and subgroups, internally operated quality control processes, external evaluations of outputs, outcomes;
- Internal and external auditing requirements, procedures, and use of results; and
- External oversight by executive, legislative, and judicial bodies as well as direct public scrutiny.

98. Doubtless other examples could be added to these lists. Incorporating such functions into an overall assessment of Integrity Framework has several key aspects:

- **Enumerating the functions:** which sorts of controls and processes does an agency claim to have in place, and which does it lack?
- **Verifying performance:** of the controls and evaluations an agency claims to have in place, which are (or are not) performed on a regular basis?
- **Follow-through:** for controls and evaluations that are performed on a regular basis, who has access to results, who has responsibility for using results to produce needed changes, and are those changes made?
- **Closing the loop:** do changes made in response to internal integrity controls produce improvements in behaviour and performance?

99. All four steps are essential if analysts are to be confident that formal integrity measures and controls amount to more than just a list of good ideas.

100. Complementing the formal measures listed above, and helping verify their effectiveness as well, are **informal** aspects of the Integrity Framework. Some relate to the **working climate** within the agency:

- Do officials understand what integrity means, both generally and in detail, and do they take it seriously?
- Do they see existing controls as credible, feasible, and effective?
- Or do they, by contrast, witness frequent breaches of integrity despite (or perhaps because of) the ways those controls are set up and implemented?

101. Other informal aspects relate to the inner context of public organisations – an important element of the Integrity Framework – that includes management instruments, status, job satisfaction, and similar issues that might become threats to integrity at the individual level:

- Do officials believe their superiors, and/or private clients, tolerate or expect breaches of integrity?
- Do officials believe they will be rewarded or punished for reporting misconduct or other integrity violations?
- Do officials *know* about channels for reporting integrity problems, such as “hotlines”, inspectors general, or integrity officials and tribunals?
- Do officials believe they would have to appeal to outside parties (the press, NGOs, political leaders) in order to address integrity violations in their working environment?
- Do officials see their work, and the goals of their agencies, as clear and worthwhile?
- Do officials believe their jobs give them sufficient status, fulfilment and income?

- Do officials perceive citizens and political leaders as indifferent to their efforts, or even as adversaries creating an “us-versus-them” fashion?
- Do officials *trust* co-workers, superiors, political leaders, and citizens?
- Are officials able to use their discretion, when needed, to accomplish worthwhile goals? Do they frequently witness in which such discretion is abused?

102. Again, many other items could be added to the list. As the questions above suggest, a key strategy for assessing informal aspects and the context of the Integrity Framework will be surveys of officials. A variety of examples of existing data are available: the Ethics Resource Center, in the United States, conducts regular Ethics Surveys in the government, business and non-profit sectors.<sup>26</sup> Moreover, the Ethics Resource Center recently conducted an integrity survey in the New Zealand public service<sup>27</sup>. The Independent Commission Against Corruption (ICAC) of New South Wales, Australia, has developed a sophisticated survey-based approach for assessing corruption risks in the government sector of that state (Gorta, 2006).

103. It might be argued that surveys of these sorts just reintroduce the same perception-based approaches that have already been developed, and that have been shown above to be quite problematical. But such is not necessarily the case. The sorts of surveys proposed here are **qualitative**, not perception-based. They are not intended to assign numerical corruption or integrity rankings to agencies or officials, and do not simply ask officials “how corrupt is agency X”. Instead, the focus is on officials’ personal knowledge and experience: how often do they perceive misconduct, how do they experience agency controls, how do they understand their roles and the agendas of their agencies, and so forth. As a consequence the results are not nearly as subjective as are perception-based surveys. Similarly, across-the-board corruption or integrity ratings are not the point; a strong point of the New South Wales approach, for example, is its focus on the identification of corruption risks, and on building resistance to corruption (Gorta, 2006). Where officials do not see their jobs as rewarding, believe integrity rules and controls are ineffective, witness frequent examples of misconduct, and believe they are better off not reporting such cases, vulnerability to corruption and related threats to integrity are severe. Where such attitudes are in place, formal elements of the Integrity Framework are likely to be ineffective, or even to serve as smokescreens for corruption and other problems. Even if major resources and top-level commitment are being devoted to those formal measures, controls that are not *seen* to be effective at the level of officials’ own work and experience must be re-examined. As with the input, process and output data in the previous section, successful revisions of formal controls can be verified by changes in the responses of officials.

104. To be sure, such surveys are detailed and, at least in the first few rounds of administration, might be relatively expensive. They must be conducted in a way that makes it clear to officials that they will not suffer reprisals for their responses, and that those responses – indeed, their daily experiences in the course of their work – are taken seriously. The model used by the ICAC New South Wales, in which an independent agency with a high level of credibility takes responsibility for data-gathering and analysis – offers an excellent example and useful suggestions for how such a constructive survey environment can be maintained. This and other ongoing efforts, in particular the regular ethics surveys of the ERC in the US as well as staff and users’ surveys across OECD countries, in particular in South Korea, the Netherlands and

---

<sup>26</sup> Ethics Resource Center, “Ethics World”, [www.ethicsworld.org/ethicsandemployees/nbes.php](http://www.ethicsworld.org/ethicsandemployees/nbes.php) (Viewed 8 September 2008).

<sup>27</sup> See “New Zealand State Services Integrity and Conduct Survey” at [www.ssc.govt.nz/display/document.asp?docid=6186&pagetype=content&pageno=3&displaytype=std](http://www.ssc.govt.nz/display/document.asp?docid=6186&pagetype=content&pageno=3&displaytype=std) (Viewed 9 September 2008).

New Zealand, would provide opportunity to share results and develop comparable benchmarks across jurisdictions. A cost-effective solution would be to agree and develop common questions on integrity and resistance to corruption in future survey questionnaires.

### **Benchmarking the Integrity Framework**

105. As in the corruption-risk assessment methodology discussed in the following section, data on both formal and informal elements of Integrity Framework will be most useful if they are gathered regularly over time, and are benchmarked against similar evidence from comparable agencies and jurisdictions. Here again, caution is in order in extending comparisons to the private sector; as argued earlier there are numerous contrasts between public and private organisations, both obvious and subtle, that make comparisons risky. Moreover, reflexive recommendations that government emulate business may only make matters worse, because they do not reflect sufficient appreciation of those public-private contrasts.

106. Among broadly comparable public agencies, however, comparative inventories of formal controls are quite useful. Such comparisons are similar to the “best-practices” listings popular in some quarters, and can highlight agencies and jurisdictions that have not taken minimal, reasonable, or extensive steps to implement formal controls. Similarly, with respect to many of the informal elements noted above **analysts, policy makers and managers have no way of knowing whether observed results are good or bad news without comparing them across a number of cases.** For example, 25 per cent of officials in an agency may believe they would be punished for reporting integrity issues (*e.g.* observed misconduct, such as corruption); but is that percentage high or low? Ideally policy makers and managers would hope to find that officials never witness breaches of integrity and never fear the consequences of reporting them. It is clear that if 98 per cent witness misconduct frequently then problems are severe. But in most cases policy makers and managers will not see such optimistic or dire results. Moreover, **the purpose is to identify vulnerabilities** rather than to measure how much corruption exists, or how extensive integrity is. To that end, only repeated comparisons can tell policy makers and managers whether or not the 25 per cent figure suggested above, and the range of other results policy makers and managers are actually likely to see, are or are not, reason for worry. Data that are gathered in detail, and that are benchmarked in the ways suggested here, will be “actionable”: that is, they will identify key points of vulnerability (*e.g.* lack of accessible procedures for reporting misconduct and fear of reprisal) and point directly to steps that need to be taken (*e.g.* establishing accessible reporting procedures and strengthening protection of whistle-blowers).

107. An equally important form of benchmarking is to compare results on “core” measures of the Integrity Framework to the data on supportive management measures and context discussed later. Do the absence of key formal controls, or their non-implementation on a regular basis, and do officials’ reports of frequent misconduct, correspond to long delays, numerous steps, or other “blinking indicators” of integrity problems in risk areas? Do improvements in one area, over time, show up as corresponding positive changes in others? Where such correspondence is apparent policy makers and managers can be confident in the methods, and in such integrity assessments, without attempting to measure corruption as such in any direct way.

108. Ultimately, all government activities and all efforts to assess and enhance integrity are intended to produce superior **outcomes** – not just quantitative growth in outputs or evidence of declining abuses of power, but rather the “impact and consequences of processes and outputs for the community” noted in an earlier section. Governments operating at high levels of integrity are most likely to use public power and resources competently, and to respond to citizen demands and social changes, in ways that improve the overall quality of life at individual, family, civil-society, and social levels. How can policy makers and

managers assess those outcomes, and relate them to findings on core integrity measures, supportive management and wider public sector and governance context?

### **Government at a Glance: Data on core integrity instruments**

109. For the first issue of the Government at a Glance, data on the following two areas of the Integrity Framework are proposed to be included, namely:

- **Conflict of interest** – the scope of declaration of private interests by decision makers and the level of transparency to disclose it.
- **Whistle-blowing / public interest disclosure** – procedures for whistle-blowing / public interest disclosure and protection for whistleblowers.

110. Personal bias by private interest of public officials in the decision making particularly weakens citizen's trust in government in public institutions. Standards and measures for **preventing and managing conflict of interest** are crucial to ensure that the integrity of decision making is not compromised by public officials' private interests.

111. Disclosing private interests is crucial to determine whether they could improperly influence the performance of official duties and responsibilities. Disclosures may also support the detection of illicit enrichment. This benchmark on the scope of conflict of interest shows which decision makers from the executive branch and the legislative branch of government are required to make declarations on private interest and what kind of information is included. Moreover, the benchmark shows the level of transparency indicating whether the disclosure by decision makers who are accountable to the public is available or not for public scrutiny. This dataset on conflict of interest measures focuses on elected decision makers including president, prime minister and ministers as well as members of Parliament and Congress. Data to cover other decision makers and public officials – for example the judiciary, officials working in risk areas such as public procurement – could be collected for inclusion of future edition of the GaaG.

112. The risk of corruption is significantly increased in a secretive environment. Facilitating the **reporting of misconduct** could substantially help public institutions monitor compliance and detect misconduct such as fraud and corruption. Proposed dataset addresses whether governments provide 1) accessible procedures to raise concern of violation of laws such as fraud and corruption, and 2) adequate protection of whistle-blowers. The data collected for GaaG will indicate whether countries have laws, policies or organisational rules on whistle-blowing and whether they provide procedures and protection for whistle-blowers.

## VI. PHASE II: RESISTANCE TO CORRUPTION AND RISKS: INPUT, PROCESS AND OUTPUT DATA

113. Instead of making corruption, and vulnerabilities to it, a matter of perceptions, policy makers and managers would find it more relevant to measure more directly the sorts of things governments do, and compare both current results and change over time to verifiable benchmarks. The underlying idea is to track trends in the **effects** of corruption, and in the **incentives sustaining it**, using data on selected inputs, processes, and outputs. The notion of **actionable** data requires that as much as possible collected data must directly reflect aspects of governance over which policy makers and managers have significant control – specific processes, problems, and trends about which leaders and reform advocates can do something. It follows from that notion that such data must track changes in sensitive and accurate ways, not only agency managers, policy makers and political leaders invaluable feedback on the effects of their efforts, but also – as noted – enabling them to take credit for success or progress, or fixing accountability for failure.

### Basic examples

114. Consider, as an example of data important to both good governance and vulnerabilities to corruption, bureaucratic delay. Assume that in City A getting a building permit involves 33 steps and takes seven weeks, while in City B the same process typically involves 4 steps and takes three days. Policy makers and managers cannot measure corruption in those two agencies directly, but theory and anecdotal evidence suggest that the numerous steps and long delays in City A are at least partly **effects** of corruption: bureaucrats have found they can make money by contriving new requirements and delays. Even more clearly, they are **incentives** sustaining it, and creating vulnerabilities: construction firms, for example, knowing that time is money and facing numerous bureaucratic “toll gates” and long delays, find it quicker and cheaper to pay up. In City B, by contrast, officials have fewer opportunities to demand payments and contractors have fewer time-related incentives to bribe. Quite likely there is less corruption in City B’s agency: even if bureaucrats responsible for one of the four steps demand payments, resulting snarls and delays are likely to become visible, and intolerable, more quickly in an otherwise rapid process. Further, if City A reduces the number of steps and amount of time involved in its process policy makers and managers can infer that opportunities to corruption and (with caution) that corruption will be reduced there as well. A relatively<sup>28</sup> fast government process, other things being equal, is preferable to a slow one; accomplishing a comparable task in fewer steps is a powerful signal that tampering with the process is less common and, in all likelihood, less tolerated. While it is possible that the faster process could be of lesser quality, with substandard construction projects slipping through a faster process for lack of scrutiny, that risk can also be assessed through examination of outcomes: independent inspections of randomly-selected projects, keeping track of citizens’ and others’ complaints, and other possibilities to be discussed in a later section.

115. Information on the number of steps and time required for such routine government functions can be gathered relatively easily on a regular basis<sup>29</sup>. The results are both easily understood and **verifiable**: an agency can compile them from its own records, while investigators and citizen groups might file applications of their own, sending individuals through various administrative processes in order to gather their own data. Results can be published regularly as part of a “governance barometer”, and can be

---

<sup>28</sup> “Fast” is obviously a relative term: taking ten-weeks to issue a passport is extremely slow, poor performance, while approving a proposal to build a nuclear power plant in only ten weeks would for most people qualify as extreme (and unwise) haste.

<sup>29</sup> For example, indicators on red tape by OECD and WB’s doing business indicators.

compared over time and among jurisdictions. The gathering of data can be targeted upon specific agencies, public functions, and locations, and can be used in pre-reform phases to identify priority targets.

116. The data could be easily gathered, clearly linked to the effects of corruption and to its sustaining incentives, comparable from one place and time to another, and easily understood by citizens and non-specialists. A few examples might include:

- The frequency and trends in exceptions, for example using non-competitive tendering, and discretionary purchases, in public procurement.
- Modifications of public contracts after awarding in the public procurement process, speed and accuracy with which vendors' invoices are paid.
- Time, expense, and the number of steps involved in establishing a small business or a corporation.
- Time and charges for specific services such as passports, licenses, and permits.
- Time and charges involved in obtaining routine information, copies of documents.
- Time, number of steps, and frequency and range of variations involved in tax assessments and other revenue collections.
- Trends in the numbers of licenses, permits, and in subsidy or benefit payments, granted by a given agency.
- Whether purchases are delivered in the quantities, quality, and places intended, and in a timely fashion.

117. There is a consistent logic underlying these and other such data: if one unit of government pays fifty per cent more for petrol than another; charges unusually high (or low) royalties for extracting natural resources; needs ten employees to accomplish tasks performed elsewhere by four; publishes budgets that are routinely far off the mark in terms of actual revenues, expenditures, and payrolls; has comparatively high rates of zoning and regulatory variances, or of tax exemptions and non-collection; conducts far more (or many fewer) regulatory inspections than other comparable units of government; or routinely receives poor marks for its public services from citizens or businesses, such facts point again to the effects of corruption and a range of incentives and gaps facilitating further abuses. Trends toward, or surpassing, benchmarks established by comparisons to other governments, or to open markets – cannot tell us directly that corruption is in decline, but do suggest that opportunities and incentives for it are being reduced, and that its effects (often closely related in practice to the incentives) are being reduced.

118. Third parties can play useful roles in data-gathering. Citizens or advocacy groups might submit applications for information and documents on a regular basis, keeping records on the time, expense, and difficulties involved. Business and trade associations could also be enlisted as participants in assessments of the latter sort. Businesses involved in routine commercial transactions with government can, if persuaded, provide critical data: the Extractive Industries Transparency Initiative (EITI),<sup>30</sup> launched in 2002, seeks to get extractive industries (oil, uranium, and many other resources), and the governments with which they deal, to publish the prices actually paid and received for extracted commodities. Prices that are

---

<sup>30</sup> Extractive Industries Transparency Initiative (EITI); Online at <http://eitransparency.org/> (Viewed 8 September 2008).

out of line with market benchmarks – either too high or too low – may be signs of corruption and other integrity problems. Twenty-two “candidate countries” have committed to implement EITI principles, but no country is as yet “EITI Compliant”. EITI receives further support from a 300-member “Publish What You Pay” coalition of NGOs.<sup>31</sup>

119. A far riskier third-party tactic borrows “secret shopper” techniques from the popular media, sending “clients” out to public agencies to conduct transactions. Unbeknownst to officials, such “clients” are equipped with hidden cameras and recording devices to capture evidence of demands for payments and other abuses. While such integrity testing tactics have been used effectively in Mexico and elsewhere to document specific allegations of corruption, they should be avoided in any comprehensive integrity assessment strategy, as they will likely lead officials to withhold data and co-operation, and can also undermine trust. To be useful over the long run, third-party involvement should be open and transparent.

### Compared to what? Building benchmarks

120. By itself a given datum tells us little about whether the effects of past corruption and incentives sustaining it are substantial or minimal, and nothing at all about trends. Data need to be assessed against **benchmarks**. Benchmarks are expressed, or anchored, in terms of verifiable evidence external to the phenomenon being measured. Metaphorically, the term derives from the practice of a tailor, optician, or laboratory scientist who might literally make marks on the bench where the work is being done in order to compare dimensions essential to the work at hand. The marks can be precise, do not depend upon individual perceptions, stay in the same place from one day to the next, and provide a reliable standard allowing comparisons **in absolute terms**: the width of a piece of cloth, the focal length of a lens, or the angle and length of an inclined plane.

121. For assessing integrity benchmarking refers to standards against which governments or agencies can be measured or judged – standards that should be defined in terms of positive values and desired outcomes. Such comparisons should extend over time. Examples might include the amount of time, and number of steps, required to obtain a building permit across jurisdictions. Similarly, each of the other examples of data noted in the previous section can also be thought of as a benchmark, once data have been assembled from several comparable cases. The OECD is already frequently approached for benchmarking data on a variety of issues (Manning, Kraan, and Malinska, OECD 2006: 12) – a trend that will only intensify with the appearance of the *Government at a Glance* series – and arguably possesses critical advantages in this area in terms of data holdings and expertise.

122. “Benchmarks” are both measurements and goals expressed in terms of some positive value – speed, for example, the efficient use of public funds, or high-quality public services. Examples, respectively, might be the proportion of exceptions to competitive tendering in public procurement, percentage of modifications of awarded contracts, late payments, the number of steps and amount of time involved (*e.g.* in obtaining a license or permit), the price for goods and services, and levels of citizen satisfaction (*e.g.* policing). For more discussion of such data, and of the ways in which they reflect the effects of corruption and the incentives sustaining it, see the table in Annex 1, below. As such they do not assess the overall state of the grand problem of corruption, but rather state positive and, it is hoped, attainable goals.

123. Benchmarks should not be confused with “targets”, which once attained have a way of becoming ceilings. Instead, they embody comparative standards based on a range of observations – other governments or agencies, for example, or in selected cases comparable practices in the private sector – that will change in ways reflecting established practices. For many of the benchmarks proposed in Annex 1, for

---

<sup>31</sup> Publish What You Pay; Online at [www.publishwhatyoupay.org/english/](http://www.publishwhatyoupay.org/english/) (Viewed 8 September 2008).

example, improvements across several jurisdictions, once put into motion, will produce gradually increasing benchmarks that provide incentives for still further improvements. Declines can also be made apparent.

124. Benchmarks are also specific – not summary data of governance across whole societies or jurisdiction, but evidence on how quickly an agency does its work or how far governments depart from market levels in the prices they pay for basic commodities. Thus multiple benchmarks are not only possible but desirable. Data from several agencies or jurisdictions, gathered over time, allow us to watch for very low numbers (a deliberately sluggish process designed to extract bribes), very high figures (frequent inspections as a form of bureaucratic or political harassment), or “spikes” or rapid declines in prices paid, percentages of taxes subject to exceptions, and so forth.

125. Properly and consistently compiled, benchmarks can be unobtrusive – compiled at relatively low cost and with little or no disruption of agency operations – and can be validated by citizens and others outside the agency itself. Benchmarks, because of their detail and, if desired, sensitivity to changing conditions, do not lend themselves to league-table rankings. But they can capture something of the real complexity of both governing and reform, and are likely to reveal actual rather than perceived changes. Most important of all, benchmarks and data can and should be **actionable** – that is, should directly measure processes, incentives and risks that can be changed, and for which accountability can be clearly specified, and should tell us which specific agencies or programs are most in need of attention. To return to the first example above, where delays are extensive and steps are numerous in a licensing or permit process, streamlining those processes is both called for and is a way of assessing subsequent progress.

126. While “benchmarking” is also a form of measurement, and thus subject to all of the **caveats** noted above, the idea differs from the notion of measuring corruption by various indices in ways that are subtle but important. Corruption or governance indices are not only subject to definitional vagueness and major questions of validity; they are also, of necessity, **relative rankings**: units of analysis may be rated in comparison to each other, but where the phenomenon in question cannot be directly measured and (thus) policy makers and managers lack a clear-cut zero point – as will always be the case with corruption – such rankings are not anchored in anything external to the scores themselves. As a consequence it might be said that one agency ranks higher or lower than another in terms of some attribute, but policy makers and managers have little way of knowing how high or low it falls in an absolute sense. And as long as that is the case, tracking change over time is inherently uncertain: it is easy to imagine a country’s corruption score “improving” relative to others on a list while corruption problems for the list as a whole grow significantly worse.

127. By way of analogy, consider the rankings of universities currently popular in the United States.<sup>32</sup> “Quality” of a university can never be defined precisely, much less measured directly. It is a matter of opinions reflecting contrasting preferences and levels of information, and is influenced by a range of factors external to the institutions themselves. At best policy makers and managers can compile perceptions and a variety of secondary data (entering students’ test scores, data on financial soundness, faculty-to-student ratios) – all of which may relate to each other,<sup>33</sup> and to the concept of “quality”, in

---

<sup>32</sup> For a particularly notorious example, see the *U. S. News and World Report* “Best Colleges 2009” rankings at <http://colleges.usnews.rankingsandreviews.com/college> (Viewed 8 September 2008).

<sup>33</sup> In the university-ranking example, wealth may allow lower faculty-to-student ratios, justifying higher tuition charges, and thus reinforcing perceptions of quality in ways having nothing directly to do with the quality of education. With respect to corruption perceptions, a particularly worrisome variety of interaction among data is the “echo chamber” phenomenon, in which a country might receive high or low scores because those interviewed simply echo the views of others. For a rigorous exploration of that problem in the context of governance data, see Thomas, 2007.

complex and non-linear<sup>34</sup> ways – and merge them into a composite indicator. That latter process adds further uncertainty, as the calculations reflect a range of assumptions, preferences, and limitations inherent in the evidence. Further, how should results be reported? No institution worth ranking will have a complete absence of quality, and hence there will be no true zero point. The resulting index can then be no more than an interval-level measurement. But because so many uncertainties and assumptions are built into the analysis of underlying data, and because relationships between the data and the underlying concept of “quality” are complex and diverse, policy makers and managers cannot be sure that (say) a five-point difference is the same across all ranges of the scale. Thus policy makers and managers are left, at best, with ordinal-level data that are not anchored in any evidence beyond themselves. While one university might move up in rankings from one year to the next, policy makers and managers have no way of knowing whether it is improving in absolute terms, is simply holding its ground while others decline, or is simply declining in quality more slowly than its competitors.

128. With respect to assessing governance and corruption, the benchmarking strategy does pose some trade-offs. Because “governance”, like “corruption”, is an abstract notion with many components and manifestations, policy makers and managers are still using indirect measurements. And, as will be explained, in some societies the sorts of data policy makers and managers might want to use as benchmarks will be difficult to obtain – or, will be sensitive data **precisely because they do tell us how good a job an agency or government is doing**, or about the cumulative effects of corruption and the incentives sustaining it at any given time. The key is to choose data that are valid, reliable, tolerably precise, and clearly embody desired governance outcomes.

129. The advantages of the DB strategy, compared to relying upon corruption indices, are however decisive. By looking at data and benchmarks of actual inputs, processes, and outputs, policy makers and managers give concreteness and detail to the general idea of governance, and link assessments to positive values: integrity becomes not just the absence of corruption (unverifiable in any event) but rather a bundle of positive values and outcomes. Benchmarks are more easily interpreted, and less susceptible to disagreement about their significance, than rankings on perception scales. Best of all, in many instances it is possible to anchor and express proposed benchmarks in **externally verifiable** standards (*e.g.* percentage of exceptions to competitive tendering in public procurement, numbers of steps involved in licensing) or evidence (*e.g.* levels of citizen satisfaction). That means policy makers and managers can not only compare governments and, selectively, agencies relative to each other – if that is a goal – but can also compare them in absolute terms and over time.

130. There is nothing magic about data and benchmarking. As will be explained in sections to come, the evidence in question is often quite familiar, even mundane. Even so it may be difficult to come by, particularly in situations where benchmarking might reveal major failings. In order to be meaningful, data must be gathered, assessed, and reported on a regular basis, and the benchmarks themselves will need to be reassessed from time to time. While benchmarking strategies can point to remedies, somebody must assume the risks and responsibilities needed to make them happen. On the other hand, this strategy offers easily-understood, detailed information on governance and integrity, on sources of progress and underlying causes of problems, and on reform priorities and their effects. They offer a way to build credibility, and to forge stronger links, with civil society and reform-oriented groups in the international community. Finally, DB assessments allow successful leaders to take credit for their efforts – a major incentive toward what is often termed “political will”.

---

<sup>34</sup> Just as one example of a non-linear relationship, increasing per-student expenditures from \$10,000 to \$20,000 is likely to produce greater actual improvement in the quality of education than would increasing them from \$70,000 to \$80,000.

### Non-linear trends?

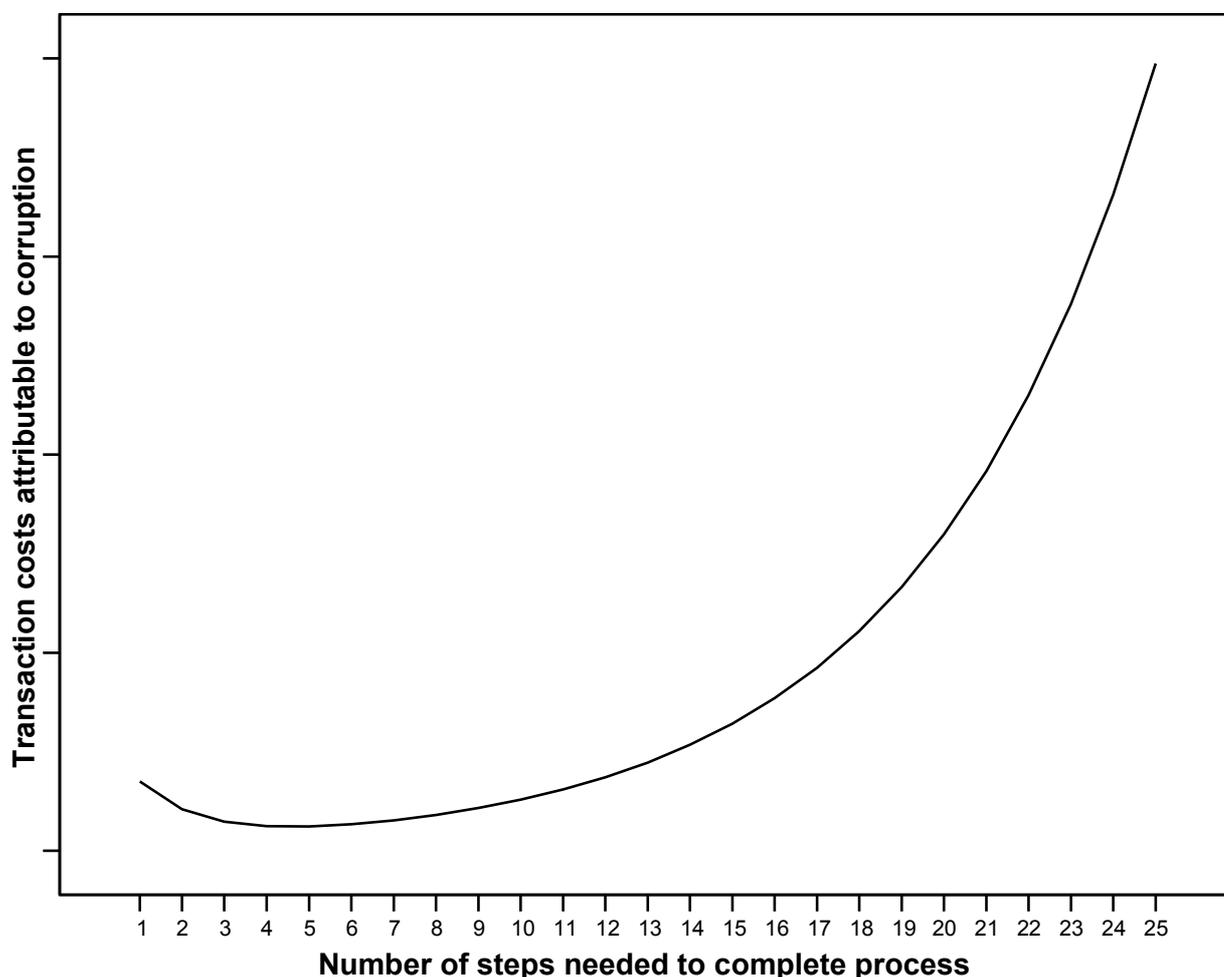
131. The DB strategy may be superior to perception-based indicators in other ways. For example, some data may not just indicate a rough rank-ordering of corruption problems – more in jurisdiction A than in B, and more in B than in C – but also hint at corruption problems that are about to become substantially worse. Similarly, some can suggest to us which government processes might be operating at near-optimal levels, and thus not be urgent priorities when it comes to allocating anti-corruption resources.

132. Consider the role of monopolies, either contrived or unintended, in bureaucratic processes. Klitgaard (1988: 75), among others, has indicated the central role of monopolies in corruption: an official who has exclusive control over access to a benefit (or over ways to avoid a tax, regulation or punishment) possesses a valuable resource, open under many conditions to corrupt exploitation. Further, Shleifer and Vishny have made a key distinction between **joint** and **independent monopolies**, illustrated by the example of toll gates along a highway, or the barriers historically erected along waterways, at which payments must be made in order to pass. In a joint monopoly situation one payment buys passage along the entire route; a somewhat analogous situation might be one in which monopolists co-ordinate tolls. Where there is a single toll to pay, or where joint monopolists co-ordinate effectively, there are incentives to keep tolls low enough to sustain a profitable flow of traffic, and to bring anyone who deviates from such arrangements back into line. Independent monopolists, by contrast, are likely to push their prices as high as they possibly can and, as monopolists, face no constraints. Tolls, or transaction costs, are likely to become so high that traffic will shift to other routes or simply cease (Shleifer and Vishny, 1993: 608).

133. By analogy, a given bureaucratic process may offer numerous monopolies. “Squeeze points” may in fact proliferate: there can be many stages in a process at which an official can put pressure upon a citizen for illicit payments, and/or where the incentive to pay up can be quite high because the official’s decision may make the difference between success or failure. Moreover, the more stages there are in an administrative process the more difficult it becomes to maintain accountability – and, where accountability is weak officials will have many opportunities to arrange for new requirements and delays, and few reasons not to. Officials may become quite skilful at contriving new monopolies, or “toll gates” that have to be passed: one may devise a lengthy new form that may have to be completed, can be arbitrarily rejected – or can be waived, for a price. Another may develop a bad habit of “losing” paperwork until the citizen pays up; and so forth.

134. At one point it was commonly said that corruption could be “functional”, breaking through bottlenecks and greasing the wheels of bureaucracy (Leff, 1964; Bayley, 1966). But such arguments overlooked a critical fact: there is no finite amount of “inefficiency”. Instead, as policy makers and managers have seen officials – particularly where they are underpaid, alienated, or relegated to low status – can contrive any amount of inefficiency if doing so seems likely to be profitable. As the number of steps in a bureaucratic process rise, co-ordination among monopolists becomes harder to enforce. As rising costs drive “traffic” – applicants or clients – out of the system, free-agent corrupt monopolists may raise prices still higher to compensate for lost income. As a result transaction costs – not just the material price of bribes, but other sorts of disruptions, delays, and uncertainties – seem likely to increase at an increasing rate. Figure 2 illustrates this idea in a purely schematic way:

Figure 1. Trends in transaction costs as administrative steps proliferate



135. At some point, a growing number of steps might be expected to cause the entire process to break down; whether matters would then lapse into gridlock or chaos is hard to say. The graphic approximation also features a slight upward turn in the curve as the number of steps falls toward one. An official controlling a single essential step operates an independent monopoly; moreover a venal private client who has reason to think a single bribe will directly produce results might have added temptation to engage in “one-stop shopping”. Policy makers and managers would still not expect costs to rise as high as they do on the right, however, for one or two payments in effect buy movement through the entire bureaucratic passage and, from a practical standpoint, accountability might be easier to maintain where only a very small number of steps are involved. (The number of steps along the X axis above is arbitrary: only with data and benchmarks can policy makers and managers guess whether an actual turning point is at about fifteen steps, as above, at seven, or at thirty.)

136. A similar logic applies to the **time** required to complete a process. The more slowly matters move along, the greater the incentive to resort to “speed money” payments. Expeditious processes are less likely to be marked by extensive corruption. But at the very fast end of the continuum incentives to corruption might rise again, as clients fearing unfavourable action, or that competitors might obtain very quick positive outcomes for the right price, seek to “buy time”. Just how slow is “slow” is a complex matter, one

that will vary from one kind of process to the next: a six-week wait for a driving license would strike most as outrageous, while a three-month process for licensing an electric power plant would be far too short.

137. The examples above become more complicated in practice. Whether or not there are multiple parallel processes, or a single one, and whether or not clients can choose who handles a given case, will affect incentives and opportunities for co-ordination among officials and the behaviour of clients. (An example would be an office at which a given process – say, obtaining a driving license – can be performed at any of three or four windows at which one waits in separate queues, versus an office where there are multiple windows but just a single queue, versus a third in which all must queue for access to a single window.) Where market or market-like processes have been put in place, perhaps by various kinds of auction procedures, other complexities arise (Andvig, 1995; Boehm and Olaya, 2006). Still, the basic point of the graph above is that the DB strategy can illuminate the effects of past corruption, and current vulnerabilities to it, in subtle ways.

138. Klitgaard (1988) also offers the interesting argument that the optimal level of corruption in an agency or process will often be greater than zero. His logic is straightforward: while the benefits of a marginal amount of corruption control will be high at the outset they are likely to fall, while the marginal costs of further controls will probably rise. At some point the curves cross, and the costs of additional corruption controls exceed the benefits they will produce; further reductions in the number of bureaucratic steps, for example, may be counterproductive by leading to hasty and flawed decisions. Similarly, other kinds of controls – writing more rules for officials to follow, more intrusive supervision, excessive investigation into officials' private lives and finances – can have real costs, not just via implementation but also in terms of reduced output, lower morale, greater inflexibility, discouraging innovation, or the loss of skilled employees to less oppressive work environments (Anechiarico and Jacobs, 1996).

139. This dimension of costs relates to the DB strategy in at least two ways: first, gathering government data need not, and definitely should not, be particularly costly or intrusive. Second, any effort to assess which controls *are* costly and intrusive, or to estimate the point at which further reform would be more costly than beneficial, will be far more precise if reliable data are collected, and judged against appropriate benchmarks of outputs and outcomes, on a frequent basis.

### Specific kinds of benchmarks

140. Benchmarks can come in several forms, and may be drawn not only from agency data but also from a variety of surveys (Reinikka and Svensson, 2003; Reinikka and Svensson, 2006). Some possible examples:

- **Statistical norms based on multiple government organisations** – If the frequency to use exceptions to competitive tendering in public procurement, or the rate of modification of awarded contracts, or the amount of time required to issue a license, or the amount paid for basic commodities such as fuel or basic services, is significantly greater or less in one public organisation or jurisdiction than the median for such data across a number of comparable public organisations or jurisdictions, that organisation or jurisdiction may have significant vulnerabilities to and incentives for corruption (see, for example, Golden and Picci, 2005; Olken, 2005).
- **Norms in private-sector organisations** – Some administrative functions – though by no means all – may usefully be compared to private-sector benchmarks. Payment of invoices or procurement might be examples, although other functions (*e.g.* distribution of social security benefits, or licensing) may quite rightly be subject to more extensive oversight, or may have no direct private-sector counterpart, making some comparisons risky or misleading.

- **Statistical norms based on market transactions** – Similarly, some kinds of procurement can be compared to private-sector prices and procedures, although here again public agencies are often subject to restrictions that businesses do not face. Fees are a more complicated question, as public agencies generally do not sell commodities in a competitive market; still, charges for publicly provided goods and services may be comparable to private alternatives, as may the shares of revenues consumed by administrative overhead.
- **Expert assessments** – Where comparable markets or private-sector processes do not exist (as in a society that has only recently or partially liberalised its economy), or where a jurisdiction, agency, or task is essentially unique (a city that is by far the largest in a country, or an agency charged with granting exceptions to environmental regulations), expert panels might allow us to define reasonable standards. Expert assessments are potentially useful for all four of the major forms of data – inputs, processes, outputs, and outcomes.
- **Statistical projections of expected outputs** – Even where directly comparable examples cannot be found, statistical models might yield useful benchmarks. Road-building in a mountainous jurisdiction, or policing in an exceptionally densely-populated urban area, might be inherently more expensive than similar functions performed elsewhere, but multivariate statistical models can be created (and over time, refined) allowing public officials to estimate a reasonable level or range of costs (see, as a possible example, Dincer, Ellis, and Waddell, 2006).
- **Comparisons of processes and outputs as a function of agency resources** – Levels and trends in processes and outputs can and should be judged in the context of resources (funding, staffing levels, and so forth) that are available. An agency whose data improve by five per cent might seem to have registered significant improvement until policy makers and managers judge such outcomes against a ten per cent increase in resources, for example. Another whose services are slightly less satisfying to citizens than those of its counterparts elsewhere might still be well-governed if it achieves such results with a workforce that is twenty per cent smaller than the norm elsewhere. That agency, unlike the first example, might be a good bet to use an influx of new resources effectively.
- **Comparisons of substantive outcomes as a function of agency activities** – Is a health department in one city that conducts significantly more restaurant inspections than a multi-city norm engaged in shakedown operations? Is one that conducts many fewer inspections bought off, or engaging in favouritism? One clue might be the actual level of cleanliness, or frequency of food-borne illnesses, as assessed by citizen surveys, independent data-gathering, and public health statistics.

141. Any and all such data and benchmarks should be gathered **repeatedly** in order to establish trends.

142. A variety of benchmarking strategies is needed because of problems and complexities inherent in the DB strategy. Some societies may have a large number of broadly comparable local jurisdictions and a vigorous, open market; others may be dominated by one large city or have economies in which prevailing price mechanisms are poorly integrated. In some places officials may be receptive to the idea of compiling and publishing data, and may even take the lead; in others they will resist efforts to gather data or may simply falsify the numbers. Comparisons to private-sector data (which may themselves be scarce or of dubious reliability) should be made with caution, as government is emphatically not just another business organisation. Some benchmarks, such as a multi-agency and city mean or median figure for prices paid for same good and services, will be affected by the compliance or resistance of officials and managers, and may thus need to be compared to external standards. For example, where agencies and cities are monitoring and upgrading procurement practices, the benchmark figure may move toward market prices,

providing a meaningful “moving target” that agencies can shoot for. On the other hand, an agency or city that performs better than the multi-agency or city median may not be accomplishing much if that median itself is moving away from the market norm.

## VII. THE OUTER CONTEXT: OUTCOMES

143. The notion of outcomes is deliberately broadly defined, emphasising the basic reasons why integrity in government is essential: service to the community and a better life for its citizens, resulting in trust in government and public services which is a key factor in the wider context of society and public life. Outcomes are distinct from “outputs” in several senses, beginning with its normative orientation: the issue is not just how much of something government does, using what sorts of inputs, but rather whether it does good things that reflect the needs, wishes, and values of citizens. While the idea of outcomes is, for current purposes, neutral with respect to the institutional arrangements required to govern in such fashion, and with respect to ideological and philosophical outlooks on what “good outcomes” are, it is not at all neutral with respect to the basic idea of government as serving the people and being guided, in important respects, by popular choices and preferences.

144. Outcomes are also cumulative: they include not only what is done but **how** it is done. The notion that the ends justify the means runs contrary to integrity as discussed earlier. Democratic processes, in particular transparency rules for election and party financing, participation by, and consultation with citizens, civil society groups, businesses, and other interested parties are integral to integrity, and to positive outcomes, even when they involve real costs in terms of time and resources. It is easy to imagine a government or an agency that rapidly produces outputs using modest amounts of resources, and with little or no outright corruption, that nonetheless contribute to bad outcomes; an example might be a housing authority that rapidly, and relatively inexpensively, builds housing of the wrong types in the wrong places, disregarding environmental and density issues. In that light it becomes clear that citizen participation, consultation, and evaluation of results are not luxuries but rather integral parts of governing at high levels of integrity.

145. Outcomes are often long-term considerations: the negative externalities flowing from the housing example suggested above will accumulate over many years, and will be difficult and costly to correct. Outcomes can be judged in terms of public or private goods: unwise housing decisions might overcrowd streets, parks, transport, and schools, for example, and can also exact costs at the level of private property values, families’ health care costs, and risks of crime. Many costs and benefits of both sorts will be difficult to measure and easy to disregard; often it will be particularly difficult to give public goods their proper weight in evaluating outcomes. A well-designed public transport line, for example, may bring extensive public benefits in the form of traffic that does not become congested, pollution that is not produced, and road expansions that do not have to be built. Putting monetary values on such benefits is possible, but involves major assumptions that may be difficult to explain to citizens and officials; nonetheless outcomes should be assessed as broadly as possible.

146. As the examples above make clear, measurement issues and causal connections with respect to outcomes are often complex and open to considerable debate. In many American cities, for example, mayors and chiefs of police are quick to claim credit when allegations of police corruption subside, usually attributing those outcomes to administrative and personnel reforms, and to changes in policing strategies, budgeting, and the like. But such trends – even if we could take them literal as indicating less corruption – have many causes, ranging from changes in types and amounts of criminal activity to the prosecutorial and law-enforcement priorities of state and federal agencies. Both outcome measures and favoured explanations for trends they may or may not reveal must often be viewed as provisional at best.

147. Nonetheless, assessing outcomes in broad-based, repeated, long-term ways is essential. That is so not only because outcomes are the ultimate standard of integrity, but also because the **process** of outcome assessment is valuable in itself. Such assessments bring officials, managers, policy makers and political

leaders face-to-face with the results of their efforts – or, with the lack of results – allowing them to claim credit for success but also helping fix responsibility for failures. Both political leaders and officials will be encouraged to look not only at budgets and processes, but at the “big picture” of how well government is serving the community.

148. Adequate assessment processes bring citizens into the process as well, regularly soliciting evaluations from them which then become the basis for public-private consultation about the best steps forward. The results of such processes can not only guide improvements to processes and substantive agendas; they can also help individual officials connect their work to real results for real people, while making citizens equally aware of officials and the challenges they face.

149. Citizen evaluations and consultation processes must not be allowed to become ritualised – a regular exercise carried out for the sake of appearances, or only to meet some formal mandate. Similarly, the “citizen side” of such processes cannot become monopolised by a few organised groups or their leaders; openness is a must for participation as well. And at least some components of outcome assessment must be aimed at general quality-of-life issues, and at trust in and satisfaction with government, rather than being tied to specific agencies, functions, and programmes. To do only the latter sorts of outcome assessment may contribute to freezing current policy agendas in place while making government less aware of new problems and emergent priorities.

150. A final potential problem is a familiar one for which no obvious solution exists, but must be kept in mind nonetheless. Social values, preferences and tastes will not be homogeneous. Some citizens may value clean air more highly than some of their neighbours, who hold schools as their top priorities. Further, a valued outcome for some may be a negative one for others: holding land open for green space may be seen in very negative terms by those who need affordable housing, or – more of a blunt example – reductions in poverty may be viewed by businesspeople as driving up the cost of labour. In one sense these sorts of dilemmas are why political processes and leaders are valuable, and should not be seen as just noise and interference in the administrative process: skilled political leadership can broker compromises across such lines of conflict, or build a base of support sufficient to mobilise solid majorities behind certain choices. Even then, outcomes assessments can help frame such choices in clearer terms for political leaders and administrative officials alike, and could conceivably educate some citizen leaders and civil society groups in the trade-offs and limitations that are inevitable parts of governing. In any event, outcome assessments should not be used to resolve such differences, or to impose any uniform set of preferences upon the community. Knowledge of conflicting priorities and preferences is valuable to all in government; repeated assessments over time can help policy makers and managers see how integrity measures might ease such trade-offs through improved outcomes and responsiveness of government.

### **Varieties of outcomes**

151. Outcomes are as diverse as the societies within which they are measured, and this paper cannot develop any comprehensive inventory of them. Some indicators may be “hard” statistical evidence on quality of services, or survey data on citizen satisfaction with the performance of various agencies. But more subtle outcomes shaped by levels of integrity and accountability are critical as well:

- **Honesty:** Do citizens believe public employees and agencies conduct themselves truthfully, handle funds, resources, and information with care, and are frank with the public regarding negative as well as positive outcomes and aspects of policy?
- **Transparency:** Does the public have a reasonable degree of access to information, and to decision making processes, within government?

- **Responsiveness:** Do officials and agencies view citizens and their values as central to their duties, rather than as burdens or in “us-versus-them” adversarial terms?
- **Redress and appeal:** Can citizens request that decisions and actions be reviewed, and where evidence warrants it, be reversed? Can flawed policy be opened up for reconsideration at citizen initiative, whether through political processes, internal review mechanisms, or both?

152. To bring the outcomes domain into the assessments of integrity on a systematic basis, and to be confident that any observed changes are actually consequences of the steps they have taken, policy makers and managers will need to compare – and, to **emphasise** the value of comparing – outcomes data to integrity data across domains, again creating opportunities for agencies and officials to claim credit for progress. Agencies that contend outcome measures do not adequately demonstrate the effects of their activities would be welcome to devise measures of their own, which could easily be incorporated into overall integrity assessments.

153. The goal, to be discussed in greater detail in a section below, is to explore the relationship among levels and trends in vulnerabilities to corruption, in formal and informal aspects of the Integrity Framework, and data on outcomes. Analysts will not usually be able to demonstrate causality, by strict criteria, among changes in those three domains, but in part because of benchmarking each family of data can be used to test the validity of the others. Where corruption risks are low or being reduced, formal aspects of the Integrity Framework are in place and are being applied, informal aspects of that Framework are strong or improving, and outcomes are positive or moving in positive directions, policy makers and managers can be reasonably confident that integrity levels are high and that, within a given set of general conditions, government is functioning well. Where the data are less positive, and in particular where they point in divergent directions, policy makers and managers have relatively clear indications of problems; and because the data are defined in terms of specific aspects of integrity and gathered at detailed levels, policy makers and managers will have useful indications of the sorts of actions that need to be taken in specific agencies, programs, and domains.

### **Sources of outcomes data**

154. Here too the basic approach requires that officials gather reliable data, this time at the level of whole communities or strategic segments of them, and compare them both over time and among jurisdictions. That task is eased somewhat by the fact that many kinds of data of those sorts are already being gathered by census and other authorities, NGOs, and international organisations, and by the fact that the process of gathering them does not intrude upon the inner workings of agencies in most cases.

155. On the other hand, the broad scale of data gathering, and measurement difficulties of various sorts (see, for example, the important but intangible integrity-related outcomes suggested above), do pose significant challenges. Comparisons are complex for similar reasons. Imagine a large city where citizen advocates routinely win reviews and occasional reversals of decisions, and obtain the release of important information, but where such processes are expensive and take several months. Given its urban scale, is that government failing an accountability test – or might it be succeeding?

156. For these reasons it will be necessary to emphasise to officials, political leaders and citizens alike that outcomes data and benchmarks are intended to demonstrate (and reward) progress over time, rather than to serve as hard-and-fast targets or as a basis for withholding resources. The best benchmarks may be moving averages or regularly-adjusted medians among communities of broadly similar scale and composition.

157. Third parties will be essential partners in assessing outcomes, both because they can contribute useful data and because of the value of participation and consultation. The “Citizen Report Cards” scheme for judging the quality of public services in Bangalore, India, is a well-known example (Thampi and Sekhar, 2006). Another Indian group – *Mazdoor Kisan Shakti Sangathan*, or MKSS – operating in some of the north-western states conducts broadly participatory “Social Audits” in which citizens exhaustively debate the workings and outcomes of government and vote on whether key objectives are being accomplished.<sup>35</sup>

---

<sup>35</sup> United Nations Public Administration Network (UNPAN) website <http://unpan1.un.org/intradoc/groups/public/documents/UN/UNPAN024251.pdf> (Viewed 8 September 2008); *Samar Magazine* website [www.samarmagazine.org/archive/article.php?id=128](http://www.samarmagazine.org/archive/article.php?id=128) (Viewed 8 September 2008); *India Together* website [www.indiatogether.org/stories/ncpri.htm](http://www.indiatogether.org/stories/ncpri.htm) (Viewed 8 September 2008); Parivartan website [www.parivartan.com/jansunwais.asp#What%20is%20a%20Jansunwai](http://www.parivartan.com/jansunwais.asp#What%20is%20a%20Jansunwai) (Viewed 8 September 2008).

## VIII. LINKING ASSESSMENT AND REFORM

158. The DB assessment strategy can both track the progress of reform efforts and be a good-governance activity in its own right. Developing the institutional capacity, commitment, and openness required to gather and publish such data, and to compare them openly with benchmarks, improves the management of agencies and the supervision of employees. Knowledge on the part of public officials that such data will be routinely gathered and assessed, that positive as well as negative results will have real consequences, and that assessment of their work will be based on a broader foundation of valid and reliable data, also reinforces the integrity framework. Because the strategy rests on assessment of activities already being conducted, it can be carried out without imposing extensive new administrative burdens or creating new delays; indeed, many sorts of data on core integrity instruments and their contexts are being gathered already, and the continuing development of e-governance can enable access to many other kinds.

159. DB assessments would give policy makers, political leaders and agency managers more extensive and useful feedback about government functions and programs. Publishing such data on a regular basis is a way to show the public, top executive and legislative leaders, public officials, aid donors and lenders, and potential bribe payers and recipients that reform is taking hold. Widely publicised and easily understood data offer reform leaders the opportunity to link their efforts to citizens' own problems and grievances, and to take credit (or blame) for results – a far more sustainable political basis for reform than simply calling for “political will”. Finally, targeted hard data do not depend upon perceptions and do not reduce an entire country's integrity problems to a single number. Rather, the current focus on multiple domains captures many of the dynamic relationships and processes sustaining, and threatening, integrity, and actionable data point directly to needed remedies.

160. Data that are credible, easily understood, and draw upon citizens' own experiences and participation will also be valuable ways to build civil-society input into the reform process. Where citizens believe their views on the quality of services are taken seriously and, over time, produce results, incentives to push publicly for integrity are increased while perceived risks of doing so are minimised. Perhaps best of all, such data of improving services link to citizens' own self-interests as “consumers” of services, as taxpayers, and perhaps as property and business owners. Creating and encouraging such vested interests in reform helps break down critical free-rider aspects of reform.

### **Issue of validity, reliability, precision**

161. Governance data have strengths, with respect to tracking trends in governance and corruption, that directly address some of the weaknesses in existing corruption indices, but as with all measurement strategies there are reasons for caution. Depending upon what analysts are trying to measure, contrasting issues of validity, reliability, and precision may arise:

- **Validity** of the data gathered in all domains can be kept high. Assessments of core integrity instruments and corruption risks are likely to be expressed in terms of the same sorts of functions (regulatory, service provision, administrative) and values (speed, accuracy, responsiveness) that the data and benchmarks reflect. Much the same is true of outcomes data. Assessments of core components and aspects of the Integrity Framework raise more questions, since they require evidence that formal measures are not only present but are being actively implemented and effectively functioning. Surveys assessing informal elements of the framework likewise must be designed with care so that they tap into officials' own experiences, rather than wandering off into the realm of perceptions. With respect to assessing trends in integrity and corruption, the validity involves a longer chain of reasoning, summarised by the notions that data tell us about the effects

of corruption and the incentives sustaining it. There, a key validity question will be whether changes in the data reflect changes in integrity practices in plausible ways (see the brief discussions of **construct validity** in box 1). Validity of the DB strategy still would seem substantially greater than that of relying on perception-based corruption indices.

- **Reliability** in assessing integrity raises some practical challenges. The co-operation of officials and agencies will be required on a sustained basis. Involving citizens in rating the quality of services may introduce expectations that are unrealistic or inconsistent from one time to the next. For those data, and to a lesser extent for surveys of officials as well, reliability may be difficult to establish until several cycles of assessment have been completed. But it is worth remembering that benchmarking itself helps officials judge reliability: if benchmarks gathered across numerous agencies or jurisdictions emerge as consistently plausible standards of performance, policy makers and managers can have greater confidence in the reliability of assessments. If benchmarks do not turn out to be relatively consistent over time that is a useful signal that the data-gathering process needs to be revised. Data grounded in real processes and events can be compared across time and space with much more confidence than can perceptions, and are likely to be less affected by many of the forces that can skew perceptions.
- **Precision** is also a strength of the DB strategy: not so much in the sense of adding decimal places to evaluation scores, but rather in the ways they allow officials, analysts, and reformers to zero in on particular governance problems, track trends over time, and derive prescriptions for action from the problems identified by the measures. To be sure, precision will vary from one benchmark to another: the proportion of using exception to competitive tendering or modifications of awarded contracts, the time involved to get a license can be fairly precisely recorded, while accuracy of budgeting, or of payment of invoices, will be more difficult to specify precisely. Data on citizens' preferences and informal aspects of the Integrity Framework ("core integrity instruments") may be less precise still, at least at the outset, for they measure intangible aspects of integrity, often in indirect ways. Here again, however, accumulated experience will over time enable improvements in precision where they are needed – particularly to the extent that officials are keen to demonstrate improvements – and any precision problems will also be outweighed by the ability to track trends with confidence.

### Implementing the DB strategy

162. The DB strategy will work best and remain manageable if it is implemented in phases, kept relatively simple, started in one or two comparable agencies or functions in several jurisdictions, and re-examined and revised after each iteration. The phases proposed are the followings:

- **First, formal and informal elements of the Integrity Framework in its "core part":** listing formal elements – for example conflict of interest measures including disclosure of relevant private interests of public officials – that are or are not present is a relatively straightforward task, but assessing the extent of their implementation will take more time and will raise more validity and reliability questions. Those assessments may require basic research – interviews and case studies of how integrity issues are addressed and controls are applied – in order to identify key items of data than can be gathered regularly. Noting, for example, that an agency has a conflict-of-interest policy and procedures for financial disclosures, or that it lists integrity training and retraining among its controls, is one thing; observing over time whether and how those controls are actually applied may take more time.
- **Second, corruption risks for mapping and fostering resistance:** this domain both presents with the fewest validity and reliability issues, requires the least organisational change and

investment of resources, and is the source of the “blinking indicators” that will signal the most critical vulnerabilities and threats to integrity.

- **Third, assessing output measures:** gathering data on quantifiable informal aspects of the Integrity Framework – such as data on ethical climate including level of observed misconduct in public organisations, confidence in whistle-blowing procedures and protection, etc – will take more time and will raise more validity and reliability questions. Consequently, they require very careful design and pre-testing of tailored surveys to adequately capture information on, for example, level of awareness of integrity instruments and commitment to use them.
- **Fourth, outcome measures:** as noted earlier, some general outcome measures might be readily available and others can easily be compiled from open-source data. Specific outcome data on integrity, however, will require the establishment of consultative and evaluative processes in groups of service users and the wider community, as well as careful research, using surveys and focus groups, to get at citizen values and preferences. In any event, interpreting outcomes in terms of integrity issues will not be sensible until analysts are confident that core and context data and benchmarks are well in hand.

163. Early rounds of DB activities should begin with one or two important agencies carrying out tasks that are comparable across jurisdictions. Public procurement and contract management is one example; certain regulatory and inspection functions, such as licensing and permit issuance, and public-health and fire inspections, are another. Beginning in this fashion will lead to data and benchmarks that are widely applicable and easily understood; starting with a few agencies, and **demonstrating that the DB strategy both produces useful information and highlights success**, can help produce more co-operation as the focus of assessments widens.

164. In each round it will be important, in early stages, to keep processes simple: it will be much better to have a few benchmarks in each “domain” in which analysts have confidence, and that are easily understood and acted upon, than to gather reams of data from the outset. **Reassessing both methods and results after each round of data-gathering** and benchmarking is essential; unforeseen problems and opportunities will emerge in early stages that can and should shape later ones.

165. Trust and transparency will be essential from the very beginning; political leaders, representatives of public officials, and citizens should be involved as the strategy is launched and revised, and as data are analysed. Results and their interpretation should be made widely and easily accessible, and public debate should be encouraged. Political leaders and agency managers can and should conduct that debate in positive ways, emphasising the benefits of integrity assessments and the steps that will be taken to improve results. A debate that descends into blame and fault-finding, and an assessment process that comes to be seen as a hunt for malefactors or as a tool for discrediting agencies, policies, political groupings, or government itself, will only be self-defeating. That maintaining such a positive tone is an extremely difficult challenge is very likely true, at least in some cases and in early phases; but doing so can pay real dividends to all concerned as the assessment process improves government and allows successful participants to take credit for their efforts.

### **Understanding the results**

166. Once data and benchmarks are in hand, what do they tell us? The key is to remember that integrity is assumed to be a property of, and to be sustained or undermined by, all domains of an integrity management system, as well as wider context of management and governance. Thus, there is no single dependent variable to be “explained”, and implementers of the strategy should resist the temptation to boil the data down into any one single “integrity index” or ranking (in any event, third parties will no doubt

attempt to do that for them). Instead, analysts will look for positive results and trends using three kinds of comparisons:

- **Comparisons of data from each domain to its own set of benchmarks:** without such comparisons policy makers and managers have no way of knowing whether results are good or bad, or whether problems are severe or relatively modest in scale. Further, comparisons to benchmarks are essential to judge whether, for example, processes involve too many exceptions, are too fast or too slow, steps are too many or too few, prices paid are too low, too variable, or too high, and so forth. Each of those kinds of divergence can be a signal of a distinctive threat to integrity, and requires a different sort of response.
- **Comparisons of data over time:** are results improving or deteriorating, both over time and with respect to appropriate benchmarks? Are they too variable and unpredictable? Do some integrity-building tactics have immediate benefits, while others work over the longer term? Are integrity-sustaining policies developing a measure of synergy, in the sense that success builds upon success and feedback loops (public responses to policies, for example) work in positive ways? Do some of the data need to be gathered in different ways, or to be replaced? Only with repeated assessments and careful evaluation can analysts answer any of these fundamental questions.
- **Comparisons of one domain to others:** as the full strategy is gradually implemented, comparisons among the various domains can be illuminating. Do improvements in one domain (for example, controlling discretionary decisions on exceptions to competitive tendering or modification of awarded contracts, or reducing the time and steps involved in licensing, or bringing the number and frequency of various inspections into line with patterns observed elsewhere) seem to produce positive results in others (for example, fewer integrity lapses observed among their colleagues, or greater trust and confidence on the part of service users and citizens)? Does one domain have more difficulty in approaching or surpassing its benchmarks than do others? Does feedback received from citizens, and do data on outcomes, lead to appropriate changes in the other domains? Answers to these and other questions can sharpen up the responses to integrity problems considerably, and give us a much more precise sense of whether those responses are working, than most officials, political leaders, and citizens now have. It is also worth noting that comparisons among domains are a key way to assess the validity and reliability of the data and benchmarks: if results from one domain are not consistent with what analysts are confident is happening in the others, that is an indication of which DB components need to be improved.

## IX. GOOD GOVERNANCE AND INTEGRITY: DRAWING UPON THE INTEGRITY FRAMEWORK

167. The DB strategy is most likely to be sustainable and effective if integrated into a comprehensive and coherent reform effort backed up, at all levels, by lasting interests and significant resources. For that reason any commitment to move forward with a DB strategy must take place in the context of the OECD Integrity Framework for public sector organisations. In a variety of ways, the proposed strategy and the framework are good fits.

### A comprehensive approach

168. A basic element in the OECD approach is to avoid the mistake of thinking of integrity problems as problems in isolation. While disease metaphors are quite popular – corruption is often referred to as a “cancer” – they are usually misleading. Serious corruption and low levels of integrity do not “happen to” governments or societies, but rather reflect deep-seated systemic influences (Johnston, 2005); they do not necessarily bring about the destruction of institutions, but instead may well remain entrenched for a very long time. Reducing integrity problems to the bad or negligent deeds of bad people is equally off the mark; if individual offenders are removed from government without eradicating the incentives and management problems that can be pinpointed by the DB strategy, integrity problems will be back, often in better-concealed and more sophisticated forms, and little will have changed actually.

169. If on the other hand the approach is to build a sound **integrity management system** (OECD, 2008) and to emphasise desired **outcomes** rather than inputs, procedures, and outputs in isolation the DB approach can help policy makers and managers address integrity problems at their roots. Such an approach creates three immediate challenges, outlined in the *OECD Assessment Framework* (OECD, 2005). These challenges are all significant in their own right, but it should also be remembered that they embody key virtues and advantages of the OECD approach:

- **Challenge 1:** determining what can be measured. Both the challenge and the virtue lie in the idea of assessing not corruption, but rather integrity and vulnerabilities that undermine it – including, but not limited to, risks of corruption. Clearly public officials and the public at large are against corruption; but what are they *for*, and what might success look like? The kinds of data proposed as part of the DB strategy reflect positive values: effective, accountable government processes (e.g. demonstrating control of discretionary decisions on exceptions to competitive tendering or modification of rewarded contracts) that use resources appropriately in pursuit of goals that benefit citizens and society. Further, movement toward those goals as expressed by the proposed data both reduces the effects of past threats to integrity and weakens the opportunities and incentives sustaining those threats in the present and future.
- **Challenge 2:** the *Assessment Framework* raises the issue of reliable and credible assessments, including the critical issue of **feedback**. Here, two aspects of the DB strategy are of particular relevance. First, most of the proposed data are objective, in the sense that they can be counted or measured in valid and reliable ways, and in many cases can be verified externally. Even the surveys aimed at assessing informal aspects of the Integrity Framework are grounded in officials’ experiences rather than just on opinions and perceptions, and can be validated against objective data. It is true that in some instances, such as tendering, contract management, employment, caseload, and internal budgeting data, obtaining the objective data will require the honest co-operation of agency managers, who in many instances may resist such probing at least initially. Getting around that issue is partly a matter of leadership, but also involves some important

incentives: a successful leader can take credit from improvements in DB results, while unsuccessful or venal management will become increasingly apparent the longer analysts persist in gathering data from other agencies, and the broader the base of information becomes. Where resistance is stiff or “gaming” of the data strongly suspected, data can often be gathered from clients, or from other agencies dealing with the organisation in question. Ultimately, unfavourable assessments and/or a lack of progress will require top leaders to apply pressure where it is needed. But with respect to gathering data themselves, wherever possible **voluntary co-operation** (even if on a very limited basis) is far preferable to any DB project that is seen as a threat.

- **Challenge 3:** raises the question of how to ensure impact? In effect this challenge, if met, closes a critical feedback loop. Here, the DB strategy offers three critical advantages:
  - i. **First** is that both collected data and benchmarks can be **tightly targeted** down to the level of specific government institutions, agencies, functions, and programmes. Thus, rather than working with high-level indices that, to the extent they measure anything at all, flatten out a wide range of contradictory trends, policy makers and managers can assess procurement practices in one agency (e.g. use of exceptions to competitive tendering and modification of rewarded contracts), personnel management practices in another, and the extent of field-agent discretion in a third, focusing in on quite specific problems.
  - ii. **Second**, the DB strategy’s emphasis upon verifiable, valid and reliable evidence allows policy makers and managers to assess **change** at detailed levels. Because collected data embody desirable results, including outcomes, a basic assessment of trends can be made in a preliminary way using just the data themselves: documenting improvements over time is an incentive itself. Where the collection of benchmarks has been thorough and sustained policy makers and managers can also know how close a given agency is to reasonable standards and how significant its movement has been over a given period. And where data and benchmarks are being assembled in all domains, comparisons across the Integrity Framework outlined in the previous section are possible and analysts can describe systemic aspects of sustaining integrity framework, and of threats to integrity, in dynamic terms. That in turn helps policy makers and managers understand what sorts of changes will have desired effects.
  - iii. **Third**, and perhaps most important, data and benchmarks, properly gathered and interpreted, are **actionable**: they point directly to countermeasures to be taken and then assessed for impact. If an agency is using too frequently non-competitive tendering, modifying contracts after rewarding, paying fifty per cent more than an appropriate market price for routine supplies such as stationery, or for basic services such as cleaning and maintenance of facilities, that departures are “blinking indicators” of integrity problems, *and* auditors, policy makers and managers have reason to think that bringing those processes and prices back into line with agency and market standards is a sound reform tactic. If tax assessors are issuing significantly more exemptions and exceptions than their colleagues elsewhere, policy makers and managers are likely seeing the effects of the corrupt use of discretion as well as important signals to business and citizens that bribery can reduce their tax bills; a re-examination of agent discretion and assessment procedures is indicated.

170. One qualification is important here: it would be wrong to infer that any and all departures are integrity issues alone, and have no other causes. Costs of commodities vary for legitimate reasons, as do levels of productivity, speed of bureaucratic processes, and the like. But where departures are very large auditors and analysts can (and should) investigate for the presence of such reasonable causes, and allow agency managers the opportunity to explain observed differences.

## Resource issues

171. A further advantage of the DB strategy is that it is not particularly resource-intensive, and that such resources and effort as it does require will themselves help improve governance and integrity. Gathering data is more a matter of leadership and co-operation than of money; after all, the strategy rests on assembling and disclosing information about activities that an agency is supposed to be doing already. Developing such leadership, co-operation, and the willingness to make such information public is by no means necessarily easy, but *any* governance-improvement effort that did not encourage them would fail.

172. Some investment in data-gathering capacity may be needed, but that too will help build and sustain integrity. Analysing the data, depending upon the sorts of benchmarks desired, need not be methodologically complex. As already noted, leaders and managers who are shown to be effective by collected data stand to benefit personally and politically in a variety of ways – a critical incentive in sustaining an integrity framework in public organisations. Finally, the involvement of citizens, clients and other stakeholders will not only help produce important data but is also, again, an integrity-building measure.

## The OECD's special role

173. The OECD is in a favourable position to implement a DB strategy, and the strategy itself will move the OECD's good-governance and integrity-building efforts in positive directions. The OECD has been accumulating data and experience in the governance, corruption and integrity fields for well over a decade, is frequently approached by governments and analysts for benchmarking data on a wide range of issues, and has been one of the leaders among international and inter-governmental organisations in those fields. The OECD's reach extends far beyond the core membership to include a large number of countries and regional bodies. The result is a knowledge and documentation base that has few rivals. The development of the *Government at a Glance* series reflects, and will enhance, all of those comparative advantages.

174. In addition, the DB strategy is naturally-suited to the OECD's history of promoting positive values of integrity as key pillar of good governance and public management, to its focus on building an integrity framework as opposed to simply combating specific abuses, and to its overall emphasis on desirable outcomes of governance. As public organisations refine their activities along their lines of priorities – *e.g.* emphasising more clearly defined levels of assessment, emphasising substance and outcomes as well as process, and striving for high performance' in public sector values and ethics and trust (Heintzman, 2005), the value of credible, tightly-targeted data on positive values and outcomes cannot be overstated.

## X. CASE STUDY: ASSESSING INTEGRITY IN PUBLIC PROCUREMENT

175. By way of illustrating how the DB strategy will work, this section lays out the essentials of implementing it for procurement functions over a two-year period. Procurement is both a central function at every level of government and an area with particular risks to integrity. It is also a function for which e-government advances have been rapid – see, for example, Mexico’s *Compranet*<sup>36</sup> initiatives – offering major opportunities for data collection and benchmarking.

176. Public procurement is a crucial economic activity of public sector organisations accounting for an estimated 15% of gross domestic product worldwide. The financial interests at stake and the close interaction between the public and private sectors provide enormous potential for waste and corruption and make public procurement a major risk area. Corrupt practices undermine competition in the market, with the result that the government pays an artificially high price for goods and services. Threats to integrity in the area of procurement are numerous. Kickbacks, collusion, nepotism, and unaccountable and excessive uses of discretion are clear threats. Contractors may substitute goods of lower quality than specified, in insufficient amounts (providing, say, 80 000 litres of petrol where 100 000 have been paid for), or may deliver nothing at all. Prices that are too high, compared to benchmarks from open markets or other public procurement processes, may signal corruption, but so may prices that are too low. Invoices may be over- or under-paid, be paid slowly or so rapidly that reviews are made difficult, or not paid at all. Conflicts of interest of many sorts may lead to results ranging from diversion of contracts to friends and political allies all the way up to officials who use their offices and public resources to operate thinly-disguised private businesses.

177. Although tendering, the awarding of contracts is the most regulated stage in the procurement cycle, it remains a major risk area because of its inherent complexities and crucial effect on the allocation of public resources. Public procurement is widely considered as a primary source of corruption in OECD countries and beyond. For example, bribery by international firms in OECD countries is considered more frequent in public procurement than in utilities, taxation, or the judicial system.<sup>37</sup>

178. Knowledge about risks to integrity and main sources of corruption could help decision makers and managers properly focus their efforts fostering integrity and resistance to corruption by effectively disbursing available resources. Mapping out risks to integrity and developing “red flags” or “blinking indicators” could particularly help managers with signals where close interaction between the representatives of public and private sectors take place in contract management and payment (e.g. signalling frequent modification of contracts in the execution stage, late payments, etc), or where frequent use of exceptions to competitive tendering take place.

- The OECD *Principles for Enhancing Integrity in Public Procurement*<sup>38</sup> provide a forward-looking approach to prevent risks to integrity in the entire procurement cycle, from needs assessment to contract management and payment. Developed on the basis of acknowledged good

<sup>36</sup> *Integrity in Public Procurement: Good Practice from A to Z, OECD: 2007; Compranet*: Online at [www.e-mexico.gob.mx/wb2/eMex/eMex\\_Sistema\\_Electronico\\_de\\_Contrataciones\\_Guberna](http://www.e-mexico.gob.mx/wb2/eMex/eMex_Sistema_Electronico_de_Contrataciones_Guberna) (Viewed 10 April 2008).

<sup>37</sup> The 2006 results of the World Economic Forum’s Executive Survey can be seen in *Integrity in Public Procurement: Good Practice from A to Z, OECD, 2007*. The figure shows the frequency of irregular payments connected to the top risk areas, namely utilities, taxation, procurement and judiciary.

<sup>38</sup> The OECD *Checklist for Enhancing Integrity in Public Procurement* and *Principles for Enhancing Integrity in Public Procurement* can be consulted at [www.oecd.org/ethics](http://www.oecd.org/ethics).

practices, the Principles help policy makers and managers apply good governance elements to enhance integrity in public procurement, namely enhance transparency, good management, prevention of misconduct, accountability and control. In addition, the OECD *Checklist for Enhancing Integrity in Public Procurement* provides practical guidance for procurement officials on how to implement the Principles at each stage of the procurement cycle and integrate them in daily management of public resources.

179. Procurement and public contracting have often proven quite resistant to reform. New York State enacted its 1912 “Wicks Law” (named for its legislative author), requiring most public construction projects to be divided into four prime contracts, in order to discourage “bid-shopping”: the practice by which major contractors took turns submitting uncompetitive bids for large projects.<sup>39</sup> The result, perhaps predictably, has been not only rampant inefficiency and major delays in projects, along with higher costs, but at times the rise of different sorts of collusion among contractors who simply agree (at times, with official encouragement) on how jobs should be shared out and then bid accordingly. Accountability for overall results is difficult to fix, and quality of construction suffers as a result. The Wicks Law is a perennial target for revision – efforts that may in fact bear fruit in 2008<sup>40</sup> – but over the years has been defended by powerful economic interests (a more general discussion of the Wicks Law appears in Anechiarico and Jacobs, 1996). In part the Wicks Law fell victim to the law of unintended consequences, but arguably it also failed because it was not based on a comprehensive analysis of integrity issues and incentives.

180. In light of the above mentioned aspects, a DB assessment procedure proposes to address the following major integrity issues in five aspects of the public procurement process:

- **Discretionary decision on exceptions to competitive tendering:** Ensuring open and competitive tendering is crucial to provide a level playing field for business. The frequent use of exceptions to competitive tendering may require other methods to ensure transparency and fairness in the procurement process. Competitive tendering requirements may at times be bypassed for legitimate reasons, but excessive use of discretion can signal favouritism, sub-standard performance, unqualified bidders, bid-rigging, and many other integrity problems. Statistical data on tendering procedures indicates the degree of competitive tendering used across public organisations to maximise transparency and competition in this crucial stage of the procurement process. This could draw attention to increased risks to integrity and also support the application of complementary mechanisms such as increased control and audit.
- **Conflict of interest:** Personal bias by private interest of public officials in the decision making weakens citizen’s trust in government and public institutions. Standards and measures for preventing and managing conflict of interest are crucial to ensure that the integrity of decision making is not compromised by public officials’ private interests. ‘Conflict of interest’ refers to a situation in which an official’s connections, financial interests, political loyalties, or other personal considerations could improperly influence the official duties. **Disclosing private interests** is crucial to determine whether they could improperly influence the performance of official duties and responsibilities and take necessary action for preventing or managing potential and actual conflict-of-interest situations. An official who admits to holding shares in a firm bidding on a contract, but offers the defence of doing nothing to help that firm, is nonetheless

<sup>39</sup> *New York Times*, “The Wicks Law is Ripe for Repeal” (26 December 2004). Online at <http://query.nytimes.com/gst/fullpage.html?res=9C04E0D61030F935A15751C1A9629C8B63> (Viewed 10 April 2008).

<sup>40</sup> New York State Division of the Budget, “Governor Paterson Announces Wicks Law Overhaul”: Online at [www.budget.state.ny.us/pubs/press/2008/enacted0809/enacted0809\\_wicks.html](http://www.budget.state.ny.us/pubs/press/2008/enacted0809/enacted0809_wicks.html) (Viewed 10 April 2008).

remains in a conflict of interest situation. Regular disclosures may also support the detection of illicit enrichment.

- **Project/contract management:** Public agencies and the communities they represent must get full value for their money. As contract management is a less transparent stage of the whole procurement cycle with close interaction between the representatives of public and private sectors, development of data and benchmarks – *e.g.* on modification of awarded contracts in the execution stage – could provide ‘red flags’ or ‘blinking indicators’ for managers and auditors on trends and variations. Controls and oversight must verify that goods and services received are those specified by contracts, in proper quantities and quality. Construction must use specified materials and practices; cost overruns and modifications should be kept to a reasonable and necessary minimum and justified by evolving circumstances, rather than becoming a conduit for corrupt or wasteful payments. Construction must stay as close to schedule as possible, and must be completed as specified (*e.g.* a massive ring road around Moscow was recently built with lanes slightly narrower than specified, allowing contractors to divert significant amounts of supplies and funds).
- **Invoicing and payment:** Over- or under-payment of invoices can signal diversion of funds and kickbacks. Payments that are too slow may indicate that pressures (*e.g.* personal or political) are being brought to bear on contractors, while payments that are too rapid may suggest efforts to avoid scrutiny and accountability. Slack auditing of payments may be the source and result of many kinds of integrity problems.
- **Prices:** as noted above, prices that are too high *or* too low may be ‘blinking indicators’ of kickbacks, graft or ‘skimming’ of contracts, favouritism, collusion or substandard performance. These ‘blinking indicators’ may be used in particular by external auditors.

## XI. CONCLUSION: FEASIBILITY ISSUES

181. It may be best to conclude this discussion with a few cautionary notes. While the DB strategy holds considerable promise and is not particularly resource-intensive, implementation may be far from easy – particularly in high-corruption situations, decision makers and analysts care most about. Tracking trends in corruption and reform via data and benchmarks may encounter considerable scepticism and resistance in the very agencies, and among the administrators, that are prime concerns. Top-level political backing and a judicious mix of pressure and support (e. g. from international *fora*, aid and lending agencies, etc) will be required in order to get the strategy off the ground. Strong and active civil society can also provide valuable support. A few targeted projects might be a good way to begin, particularly if they are followed up (where appropriate) by favourable incentives (such as publicity, recognition in performance evaluation resulting enhanced job security, higher salaries and status for successful administrators). Rank-and-file employees may resist as well, particularly if monitoring is seen as unduly intrusive or as an early sign of impending job cuts. There too, rewards and positive incentives for successful and effective employees will be critical, and will also reinforce broader good-government goals. A credible and skilful manager may be able to present this strategy to employees as a way to help them enhance their status and demonstrate their effectiveness to the community and to elected officials.

182. Will agency managers and political leaders allow access to data, will they “game the indicators” (Hood, 2006; on gaming of education data, see *New York Times*, 2008), and will they follow through on what the results tell them? “Gaming” can take many forms: Hood (2006) identifies three classic types – ratchet, threshold, output distortion – along with four other common variations: changing the focus of service provision, “storming” to meet targets in the short term, quietly removing a target after performance declines, and fabrication of results (Hood, 518-519). Others have organised such practices in two categories of altering output data and altering the outputs themselves.<sup>41</sup> Honest, comparable calculations are essential, and continuing consultation regarding methodology and eventual uses of data will also be essential. Benchmarks, as opposed to targets, may be *somewhat* less vulnerable to some varieties of gaming: if nothing else, some varieties would extensive collusion across jurisdictions.

183. Co-operation levels will vary, and entrenched interests in both state and society may resist benchmarking efforts. More generally, the most effective responses to these problems lie outside the benchmarking scheme itself, and will depend upon the ability of the OECD and participating organisations like it to encourage and sustain official co-operation.

184. That ability, in turn, is linked to the broader question of what benchmarking offers both individual countries and the OECD, analytically and in terms of strengthening the *Integrity Framework*. National governments, and their constituent levels and agencies, could find that DB strategy a source of enhanced credibility, not only for reform but also in terms of accountability, responsiveness and the provision of basic services. By gathering and publishing data and benchmarks on a sustained basis, and by using them to target reform efforts and assessments of outcomes, a government shows citizens, officials, the private parties with whom they deal, and possible wrongdoers too that it is serious about integrity and prepared to take targeted actions. Responsible officials within those governments – both agency heads and those charged with reform – will be able to develop a far clearer understanding of the scope and location of corruption and governance problems, of the underlying problems and incentives sustaining those difficulties, and of the effects of countermeasures. Reform and good governance, over time, can move from being mere slogans to demonstrable facts. The OECD, for its part, is an ideal organisation to pursue the

---

<sup>41</sup> OECD, “Issues in Output Measurement”, Summary point 12.

benchmarking strategy, and to expect that doing so will strengthen its already considerable advantages in the field of corruption control.

185. With respect to feasibility of DB strategy, there are positive points, namely, many of the sorts of data in question are being gathered already, and none require major methodological or technological innovations. The rapid emergence of e-government, as already noted, provides major opportunities for both data-gathering and benchmarking, and has sparked useful debate about the incentives and rewards that can sustain such innovation. Again as already noted, the DB strategy for integrity assessments can give major and visible credit for progress to successful leaders and managers. It can solidify many countries' cases for investment, aid, and technical assistance, providing verifiable and detailed evidence of positive changes, and specific diagnostic information on continuing problems, in ways that perception-based country indices can never do. If the "Opacity Index" noted above holds any validity – and quite likely it does – governments that both document their integrity situations and pursue targeted integrity policies using the DB strategy are likely to be rewarded, over time, with more favourable social context – for example increased levels of trust in public organisations and services – and economic terms, (*e.g.* lower costs for compliance). While implementing the DB strategy will require a significant amount of "front-end" effort, subsequent iterations will be less expensive and more productive as comparisons over time become available.

186. As some kinds of data, even if they can be obtained, may be difficult to analyse, the implementation of DB strategy requires **careful planning and consensus building** across jurisdictions. Such is particularly the case for procurement: some comparisons might be difficult to make and interpret (*e.g.* thresholds, prices of commodities), and considerable creativity may be required in devising data. Given the sorts of resistance and pitfalls noted above, it may well be that many of these more challenging tasks of comparison should be deferred until the compilation, publication, and tracking of simpler and more "transparent" information has had some time to take root.

187. A further problem is obvious but worth noting: just because data are actionable is no guarantee that action will be taken. A certain amount of disruption and resistance to gathering even the most basic data, and to following through on them with effective controls, are inevitable. As monitoring efforts spread, and in particular as they **begin to put pressure on sources of corruption and beneficiaries** to significant degrees, such resistance will only intensify. But there may well be a "tipping point" beyond which co-operation and improving results become more widely seen as advantageous and as enhancing one's own, and one's agency's prospects. Still, for a time at least leaders and managers may find the uncertainties of effective monitoring and reforms threatening in personal terms – or, at least, less preferable than living a quiet life – and may do little to put the information to use. Again, outside pressure and incentives may prove critical; here, again, is an area in which the OECD is well-positioned to encourage sustained commitment.

188. The DB strategy will be less suitable and effective for some public functions and corruption problems than for others. It deals with routine and repetitive functions, and as such will not tap into large, one-time "grand corruption" transactions. There are also substantive and common-sense limits as to how streamlined a process ought to be, how much an agency workforce ought to be cut, how much it ought to be monitored on an hour-by-hour basis, (Anechiarico and Jacobs, 1996), and how many inspections a staff member should be expected to perform. Some agencies will deal with qualitative or non-repetitive decisions that are less easily compared with others – planning permission applications are an example. Still, it is not necessary to gather data on all public functions where corruption might arise.

189. A key crisis point may occur when and as DB methods indicate severe integrity problems or a significant decline in integrity levels. At that point action will be required, and that may well be politically challenging. However, **three advantages of a well-implemented DB approach** should be noted:

- First, political leaders, agency managers, officials and citizens acting against integrity problems will be doing so **on the basis of sound evidence**, not hearsay or politically-driven scandal.
- Second, in most instances they will be able to **act against well-identified problems** in relatively tightly-specified target areas, rather than taking on (and possibly alienating) major segments of the entire public service.
- Third, progress resulting from appropriate efforts, and the public credit that accompanies it, can be **demonstrated clearly** and relatively quickly.

190. With most conventional reform strategies, few groups or individuals have strong incentives, other than abstract commitments to the public good, to act decisively against corruption, and relatively few clear hints as to what specifically they ought to do. With a DB process in place, by contrast, there are opportunities for opponents of corruption at many levels to claim credit for success, and the data themselves **point to remedies** that should be pursued.

191. The point, ultimately, is to create an overall climate of integrity shaped by learning; **incentives encouraging and rewarding successful processes**, outputs, and outcomes; public awareness; and positive administrative and public scrutiny in which the quality of government and service to citizens takes high priority. Over time this approach to tracking reform can help strengthen the links between citizen wellbeing and the ways in which public power and resources are used, enable those who govern effectively to be rewarded for doing so, and threats to integrity in verifiable and lasting ways.

**ANNEX 1. THE DATA AND BENCHMARK STRATEGY: POSSIBLE APPLICATIONS WITH RESPECT TO CORRUPTION**

*Process Phase*

<b>Indicator</b>	<b>Effects of past corruption</b>	<b>Incentives to further corruption</b>	<b>Possible benchmarks</b>	<b>Actions indicated</b>
<p><b>Time, steps, fees involved in registering a business:</b></p> <p>speed, costs, data/paperwork required, frequency and scope of variations, frequency of use of consultants</p>	<p><u>Too slow, elaborate, official fees too high:</u> extortion, bribery, official collusion with business competitors</p> <p><u>Too fast or variable:</u> favouritism, bribery, non-enforcement of policy</p>	<p><u>Too slow, fees too high:</u> opportunity to demand bribes and incentive to pay; access for officially-favoured competitors, “consultants”</p> <p><u>Too fast or variable:</u> favouritism, unfair competition, bid-rigging</p>	<p>Case handling in other jurisdictions:</p>	<p>Review officials’ monopoly power, discretion, accountability; enforce reasonable process standards; gather info from clients; randomly audit cases; create client advocates</p>
<p><b>Speed and accuracy with which vendors’ invoices are paid</b></p>	<p><u>Slow/inaccurate payment, underpayments:</u> extortion of vendors, official theft</p> <p><u>Very rapid payment, overpayments:</u> kickbacks, self-dealing, favouritism, “ghost vendors”, substandard goods</p>	<p><u>Departures from reasonable norms</u> allow kickbacks, theft, extortion; create cover for “ghost vendors” and self-dealing;</p>	<p>Multi-agency and private-sector norms; published process targets negotiated as appropriate with vendors</p>	<p>Real-time monitoring of procurement, payments; random auditing of tenders, contracts, provision, payment; solicit feedback from vendors</p>
<p><b>Time, steps, charges involved in obtaining routine information,</b></p> <p>passports, or documents</p>	<p><u>Too slow, elaborate, expensive:</u> extortion, bribery, lack of transparency</p> <p><u>Too fast or variable:</u> favouritism, bribery, forgery, blackmail, sale of official/confidential data</p>	<p><u>Too slow, official fees too high:</u> opportunity to demand \$, reasons to pay</p> <p><u>Too fast or variable:</u> payments to buy time, preserve confidentiality, avoid blackmail</p>	<p>Case handling in other jurisdictions: speed, fees, frequency and scope of variations</p>	<p>Review officials’ monopoly power, discretion, accountability; enforce standards for speed of handling cases; gather info from clients; random audits of cases; create client advocates</p>

Indicator	Effects of past corruption	Incentives to further corruption	Possible benchmarks	Actions indicated
<p><b>Time, steps, and frequency/scope of variations and exceptions in tax assessments, collections; and in other revenue collections</b></p>	<p><u>Too slow/elaborate/rigid:</u> extortion, bribery  <u>Too fast, or excessive variation:</u> favouritism, bribery</p>	<p><u>Too slow or too fast, excessive variations, exceptions:</u>                      Taxpayers/property owners vulnerable to manipulation, and/or believe payments can win favourable treatment</p>	<p>Case handling in other jurisdictions</p>	<p>Review assessors', collectors' monopoly power, discretion, accountability; enforce standards for handling cases; gather info from clients; random audits of cases; involve client advocates; monitor assessments</p>
<p><b>Number, frequency, quality, consistency of routine inspections of businesses; frequency of cited infractions; share of cases settled informally</b></p>	<p><u>Too many, too inconsistent:</u> extortion, often on-the-spot  <u>Too few, or low-quality:</u> bribery over the longer term; favouritism, both often involving supervisors</p>	<p><u>Too many, too variable:</u> opportunity to pressure clients; reasons to pay  <u>Too few or low-quality:</u> reward for cultivating, longer-term links with officials, political leaders</p>	<p>Similar inspection processes in other jurisdictions</p>	<p>Review internal accountability, inspectors' powers and discretion; set, enforce reasonable process targets; gather information from businesses, customers; verify results</p>
<p><b>Number of inspections performed per member of field staff in regulatory agencies</b></p>	<p><u>Too many:</u> extortion on small-medium scale  <u>Too few:</u> corrupt pressures from supervisors, bribery/extortion on larger scale; theft of time, "moonlighting"</p>	<p><u>Too many, too variable:</u> vulnerability to extortion, reasons to pay on the spot  <u>Too few:</u> signal that supervisors will reward larger/longer-term payments</p>	<p>Process data from comparable jurisdictions</p>	<p>Review internal accountability, inspectors' powers and discretion; set, enforce reasonable process targets; gather information from businesses, customers; verify inspections occur</p>

Indicator	Effects of past corruption	Incentives to further corruption	Possible benchmarks	Actions indicated
<p><b>Frequency and costs of outsourcing agency work to consultants;</b></p> <p>skill and training level workforce within an agency</p>	<p><u>Too high</u>: agency budget being diverted to politically connected “consultants”; consultants may be “middlemen” between agency and clients; agency employees may be devoting most effort elsewhere (shadow enterprises, “moonlighting”, political duties)</p>	<p><u>Too high</u>: agency budget becoming a pool of “spoils”; insufficient oversight; politically-favoured but unskilled employees hidden on payroll; politicians become “consultants”</p>	<p>Process data from comparable agencies, other jurisdictions</p>	<p>Increase oversight over outsourcing; invest in training agency workforce, weeding out the unproductive; analyze budgets and outlays in proportion to tasks accomplished</p> <p><i>(see input/output section)</i></p>

Indicator	Effects of past corruption	Incentives to further corruption	Possible benchmarks	Actions indicated
<p><b>Supervisory/staff vs line workforces</b></p>	<p><u>Too many supervisors:</u> agency has been colonised by elite factions favourites</p> <p><u>Supervisory payroll too large:</u> payroll has become an elite slush fund; funds may be diverted to shadow enterprises and corrupt deals</p> <p><u>Too many line employees:</u> payroll has become a patronage pool for unskilled political loyalists; internal accountability likely to suffer</p>	<p><u>Too many supervisors:</u> payroll open to patronage abuses</p> <p><u>Supervisory payroll too large:</u> funds available for theft, diversion</p> <p><u>Too many line employees:</u> patronage is likely; line employees likely to be underpaid, and thus to demand bribes</p>	<p>Workforce, payroll data from comparable agencies, jurisdictions; examine size payrolls in comparison to agency tasks and accomplishments</p> <p>(see <i>input/output section</i>); gather data on client satisfaction levels, where appropriate</p>	<p>Tighten auditing and internal accountability; bring proportions and size of workforces and payroll toward multi-agency norms; examine personnel processes (hiring, review) and day-to-day supervision</p>

*Input-Output Phase*

Indicator	Effects of past corruption	Incentives to further corruption	Possible benchmarks	Actions indicated
<b>Speed and accuracy with which funds received are handled</b>	<u>Slow processing</u> and/or <u>missing funds</u> : agents have established a pattern of diverting funds, either temporarily or by way of theft; <i>or</i> have been pressured to divert funds to superiors	<u>Established suboptimal processes</u> may become a <i>de facto</i> standard; those diverting significantly less, or more, money may be harassed by superiors or colleagues	<u>Data on fund handling</u> in other units of government, and in private sector; survey households and firms making payments to estimate error rates, delays	Improved auditing; establish process goals for speed, accuracy of handling payments
<b>Trends in the numbers of subsidy or benefit payments granted by an agency</b>	<p><u>Too many</u>: kickbacks, possibly shared with superiors; political pressures to distribute patronage</p> <p><u>Too few, or arbitrarily distributed</u>: favouritism, possibly in response to political pressures</p>	<p><u>Too many</u>: a signal that bribes can obtain benefits; or, of voter-buying</p> <p><u>Too few, or too arbitrarily distributed</u>: political interference; accountability, transparency problems</p>	Patterns, trends in benefit distributions in other jurisdictions	More specific legislation and implementation rules; audit successful and unsuccessful applications; review agency's internal accountability, independence; establish client advocates
<b>Prices charged for basic services (school meals, telephone equipment, etc.)</b>	<p><u>High prices</u>: graft/skimming/embezzlement</p> <p><u>Low/unusually variable prices</u>: favouritism</p>	<p><u>Extra revenues</u> finance graft</p> <p><u>Low/variable prices</u> invite favouritism, side deals between clients and officials</p>	Market prices for commodities, and median prices charged by other governments and agencies	Publish price schedules, DB results; monitor client access, quality of provision; audit revenue flows, tendering and bidding

Indicator	Effects of past corruption	Incentives to further corruption	Possible benchmarks	Actions indicated
<p><b>Prices paid for basic commodities</b> involved in a public service (e.g. fuel, food, concrete, tools)</p>	<p><u>High prices:</u> kickback schemes</p> <p><u>Unusual variation:</u> favouritism, nonstandard goods, breakdown of bidding</p>	<p><u>High prices</u> finance kickbacks</p> <p><u>Variation</u> invites exploitation of vendors, provision of substandard goods</p>	<p>Market prices for commodities, and median prices paid by other governments or agencies</p>	<p>Publish price schedules, DB results; monitor quality of provision, audit revenue flows, tendering and bidding</p>
<p><b>Quantity and quality of goods, commodities received</b></p>	<p><u>“Short” deliveries, non-deliveries, substandard or inappropriate goods:</u> skimming and diversion of funds and goods; kickback schemes</p>	<p><u>Tolerating inaccurate or non-delivery</u> shows that skimming, diversion, kickbacks can continue, and that oversight and controls are lax</p>	<p><u>Compare goods received to tenders and invoices,</u> if necessary on a random-sampling basis</p>	<p>Enhance supervision, auditing; institute prequalification with warnings, suspension, and disqualification of vendors as appropriate; publish data on vendor process</p>
<p><b>Staffing levels</b> in proportion to tasks accomplished</p>	<p><u>Too many employees:</u> political interference, nepotism, illicit enterprises run by managers convert payroll into a patronage pool</p> <p><u>Too few employees:</u> resources and workforce diverted to corrupt endeavours (<i>an unlikely finding</i>, if agency records conceal such practices)</p> <p><u>“Ghost workers”:</u> employees are paid for work elsewhere, or do not exist; payroll diverted by corrupt officials (<i>difficult to show</i> without comparing workforce records to actual tasks and accomplishments)</p>	<p><u>Too many:</u> weak auditing, accountability allow, conceal payroll abuse; political interference</p> <p><u>Too few:</u> weak auditing, accountability allow payroll and resources to be diverted</p> <p><u>“Ghost workers”:</u> extremely weak controls allow theft of payroll; pervasive corruption likely sustained by involvement of top leaders and managers</p>	<p>Workforce, payroll data and data processes from similar agencies in other jurisdictions, or (<i>if appropriate</i>) to private-sector firms doing similar tasks; audit actual execution of tasks.”Ghost worker” findings are <u>likely to emerge as secondary findings</u> based on analysis of other indicators and benchmarks</p>	<p>Tighten external oversight, internal auditing, and managerial/workforce supervision; monitor “output” and “outcome” indicators of tasks accomplished and client satisfaction in comparison to resources and workforce employed; revise personnel practices and data systems</p>

Indicator	Effects of past corruption	Incentives to further corruption	Possible benchmarks	Actions indicated
<p><b>Budgetary levels</b> in proportion to tasks accomplished</p>	<p><u>Too high</u>: funds are being diverted, wasted; excess or inefficient staffing reflects patronage</p> <p><u>Too low</u>: funds are apportioned among agencies based on political loyalties of managers, or their potential as reservoirs of patronage; funds are being diverted to political uses, favoured consultants</p>	<p><u>Too high</u>: budgets become political slush funds, or are diverted by managers; patronage easily hidden, workforce poorly supervised</p> <p><u>Too low</u>: funds are diverted before reaching agencies</p>	<p>Funding levels in comparable agencies in other jurisdictions, assessed in proportion to agency duties, accomplishments</p>	<p>Tighter oversight of budgetary, personnel, supervision functions</p> <p>Publish goals for agency accomplishments and adjust over time</p>

*Outcomes Phase*

<b>Indicator</b>	<b>Effects of past corruption</b>	<b>Incentives to further corruption</b>	<b>Possible benchmarks</b>	<b>Actions indicated</b>
<p><b>Amount or percentage of time, and associated costs, incurred by businesspeople in dealing with officials</b></p>	<p><u>Too much:</u> bureaucratic harassment, foot-dragging, in pursuit of bribes; extortion</p> <p><u>Too little, or wide variation:</u> favouritism; bribery initiated by business</p>	<p><u>Too much:</u> time lost, threat of fines and charges, and uncertainty increase incentive to pay</p> <p><u>Too little:</u> bribe payments can reduce administrative overhead costs, delays, uncertainty</p>	<p>Typical amounts of time spent, by sector, as established by business surveys</p>	<p>Reduce discretion; revise queuing arrangements and caseloads to break up personal links between officials and businesspeople; continued surveys of business; internal data-gathering on frequency, duration of interactions</p>

## REFERENCES

Abramo, Claudio Weber. 2007. "How Much Do Perceptions of Corruption Really Tell Us?" Economics Discussion Papers, number 2007-19 (May 4). Available at [www.economics-ejournal.org/economics/journalarticles/2008-3/?searchterm=Abramo](http://www.economics-ejournal.org/economics/journalarticles/2008-3/?searchterm=Abramo) (Viewed 4 September 2008).

Andvig, Jens Christopher. 1995. "Corruption in the North Sea Oil Industry: Issues and Assessments." *Crime, Law, and Social Change* 23:4 (December), pp. 289-313.

Anechiarico, Frank, and James B. Jacobs. 1996. *The Pursuit of Absolute Integrity*. Chicago: University of Chicago Press.

Arndt, Christiane, and Charles Oman. 2006. *Uses and Abuses of Governance Indicators*. Paris: Organisation for Economic Co-operation and Development, Development Centre Series.

Arndt, Christiane. 2008. "The Politics of Governance Ratings". Unpublished manuscript.

Babbie, Earl. 1995. *The Practice of Social Research*. Belmont, CA: Wadsworth, 7th ed.

Bayley, David. 1966. "The Effects of Corruption in a Developing Nation." *Western Political Quarterly* XIX, pp. 719-732

Boehm, Frederic, and Juanita Olaya. 2006. "Corruption in Public Contracting Auctions: The Role of Transparency in Bidding Processes". *Annals of Public and Cooperative Economics* 77:4 (November), pp. 431 – 452.

Bourgon, Jocelyne. 2007. "Responsive, Responsible, and Respected Government: Towards a New Public Administration Theory." *International Review of Administrative Sciences* 73:1, pp. 7-26.

Brown, A. J. 2006. "What are We Trying to Measure? Reviewing the basics of Corruption Definition." In Charles Sampford, Arthur Shacklock, Carmel Connors and Fredrik Galtung (eds), *Measuring Corruption* (Aldershot: Ashgate).

Cameron, Lisa A., Ananish Chaudhuri, Nisvan Erkal and Lata Gangadharan. 2005 "Do Attitudes toward Corruption Differ across Cultures? Experimental Evidence from Australia, India, Indonesia and Singapore." Available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=778464](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=778464) (Viewed 4 September 2008).

Collier, David, and S. Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research", *World Politics* 49 (April), pp. 430-451.

CBSS, 2008. Corruption Prevention in Public Administration in the countries of the Baltic Sea Region, [www.cbss.st/structure/wgdi/dbaFile17501.pdf](http://www.cbss.st/structure/wgdi/dbaFile17501.pdf).

Cushman, D. O. 1985. "Real Exchange Rate Risk, Expectations and the Level of Direct Investment." *Review of Economics and Statistics* 67:297-308.

Deichmann, Uwe, and Somik V. Lall. 2003. "Are You Satisfied? Citizen Feedback and Delivery of Urban Services." World Bank Policy Working Paper 3070, June. Online at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=636432#PaperDownload](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=636432#PaperDownload) (Viewed 4 September 2008).

Dincer, Oguzhan C., Christopher J. Ellis, and Glen R. Waddell. 2006. "Corruption, Decentralization and Yardstick Competition". Online at <http://ssrn.com/abstract=877470>. (Viewed 4 September 2008).

Di Tella, Rafael and Ernesto Schargrotsky. 2002. "Political and Economic Incentives during an Anti-Corruption Crackdown." Online at [www.utdt.edu/~eschargr/Political%20and%20Economic%20Incentives.pdf](http://www.utdt.edu/~eschargr/Political%20and%20Economic%20Incentives.pdf). (Viewed 4 September 2008).

Dobel, J.P. 1978. "The Corruption of a State." *American Political Science Review* 72:3, pp. 958-973 (September 1978).

Doig, Alan, and Stephanie McIvor. 2003. "The National Integrity System: Assessing Corruption and Reform." *Public Administration and Development* 23, 317-332.

Dornbusch, Rudiger, Daniel V. Dantas, Clarice Pechman, Roberto de Rezende Rocha and Demetrio Simoes. 1983. "The Black Market for Dollars in Brazil." NBER Reprint 364, 25-40.

Edwards, Sebastian. 1989. *Real Exchange Rates, Devaluation and Adjustment: Exchange Rate*.

Gallie, W. B. 1965. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56.

Galtung, Fredrik. 2006. "Measuring the Immeasurable: Boundaries and functions of (Macro) Corruption Indices." Ch. 6 in Sampford *et. al., op. cit.*

Gatti, Roberta, Stefano Paternostro, and Jamele Rigolini. 2003. "Individual Attitudes toward Corruption: Do Social Effects Matter?" World Bank Policy Research Paper 3122. Available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=636542](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=636542) (Viewed 4 September 2008).

Golden, Miriam A., and Lucio Picci. 2005. "Proposal for a New Measure of Corruption, Illustrated with Italian Data." *Economics & Politics* 17:1, pp. 37-75.

Gorta, Angela. 2006. "Corruption Risk Areas and Corruption Resistance." Ch. 10 (pp. 203-219) in Sampford *et. al., op. cit.*

Greene, William. 1990. *Econometric Analysis*. New York: Macmillan.

Hall, Thomas, and Glenn Yago. 2000. "Estimating the Cost of Opacity Using Sovereign Bond Spreads". Working Paper, Capital Studies Group, The Milken Institute, Santa Monica, CA 90401.

Heintzman, Ralph and Marson, Brian. 2005. "People, service and trust: is there a public sector service value chain?" *International Review of Administrative Sciences*, Vol. 71, No. 4, pp. 549-575.

Hempel, Carl G. 1952. "Fundamentals of Concept Formation in Empirical Science." *International Encyclopedia of Social Science II* (7).

Heywood, Paul. 2008. "The Politics of Perception: Use and Abuse of Transparency International's Approach to Measuring Corruption." *Political Studies* (forthcoming).

Huberts, Leo W. J. C., van den Heuvel, and Punch, 2000. "Public and Business Ethics: Similar or Different?"

Johnston, Michael. 1998. "What Can Be Done About Entrenched Corruption?", pp. 149-180 in Boris Pleskovic (ed. ), *Annual World Bank Conference on Development Economics 1997*. Washington, DC: The World Bank.

Johnston, Michael. 2005a. "Keeping the Answers, Changing the Questions: Corruption Definitions Revisited." In Ulrich von Alemann, ed. *Dimensionen politischer Korruption: Beiträge zum Stand der internationalen Forschung*. Politische Vierteljahresschrift Sonderheft 35/2005. Wiesbaden: Verlag für Sozialwissenschaften, pp. 61-76.

Johnston, Michael. 2005b. "*Es posible medir la corrupción, ¿pero podemos medir la reforma?*" ("It is Possible to Measure Corruption, But Can We Measure Reform?") *Revista Mexicana de Sociología* 2 (April-June), pp. 229-269.

Johnston, Michael. 2005c. *Syndromes of Corruption: Wealth, Power, and Democracy*. Cambridge and New York: Cambridge University Press.

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobaton. 1999. "Aggregating Governance Indicators". Washington: The World Bank; August; <http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0,,contentMDK:20789932~menuPK:1976990~pagePK:64168445~piPK:64168309~theSitePK:1740530,00.html> (Viewed 4 September 2008).

Kaufmann, D. , A. Kraay, and M. Mastruzzi. 2006. "The Worldwide Governance Indicators Project: Answering the Critics." World Bank Institute Paper (September). Available at <http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0,,contentMDK:21045419~menuPK:1976990~pagePK:64168445~piPK:64168309~theSitePK:1740530,00.html> (Viewed 4 September 2008).

Kaufmann, D, and A. Kraay. 2008. Governance Indicators: Where Are We, Where Should We Be Going?

World Bank Institute Policy Research Paper 4370. Available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1019685](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1019685) (Viewed 4 September 2008).

Kaufmann, D., A. Kraay, and M. Mastruzzi. 2008. "Governance Matters VII: Governance Indicators for 1996–2007", World Bank Institute paper; available, along with data and a variety of other sorts of documentation, at <http://info.worldbank.org/governance/wgi/resources.htm> (Viewed 5 September 2008).

Kelman, Steven. 1990. *Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance*. Washington, DC: American Enterprise Institute Press.

Kenny, Charles. 2006. "Measuring and Reducing the Impact of Corruption in Infrastructure." World Bank Working Paper 4009.

Kenny, Charles. 2006. "Measuring and Reducing the Impact of Corruption in Infrastructure." World Bank Working Paper 4099.

Available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=952071](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=952071) (Viewed 4 September 2008).

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

Knack, Stephen, and Philip Keefer. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics* 7(3): 207–27.

Knack, Stephen. 2006. "Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators." World Bank Working Paper 3968. Available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=923275](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=923275) (Viewed 4 September 2008).

Kurer, Oskar. 2005. "Corruption: An Alternative Approach to its Definition and Measurement." *Political Studies* 53:1, pp. 222–239.

Kurtz, Marcus J., and Andrew Schrank. 2007. "Growth and Governance: Models, Measures, and Mechanisms." *Journal of Politics* 69(2):538-554.

Kurtzman, Joel, Glenn Yago, and Triphon Phumiwasana. 2004. "The Global Costs of Opacity: Measuring Business and Investment Risk Worldwide." *MIT Sloan Management Review* 46:1 (October), pp. 38-44. Available at [http://galenet.galegroup.com/servlet/BCRC?vrsn=unknown&locID=nysl\\_ce\\_colgul&ste=5&docNum=A124187617](http://galenet.galegroup.com/servlet/BCRC?vrsn=unknown&locID=nysl_ce_colgul&ste=5&docNum=A124187617) (Viewed 4 September 2008).

Kurtzman, Joel, and Glenn Yago. 2008. "Opacity Index 2007-2008: Measuring Global Business Risks." Santa Monica, CA: The Milken Institute. Available at [www.kurtzgroup.com/reports/opacity\\_index\\_2.pdf](http://www.kurtzgroup.com/reports/opacity_index_2.pdf) (Viewed 4 September 2008).

Lambsdorff, Johann Graf. 2006. "Measuring Corruption – The Validity and Precision of Subjective Indicators." Ch. 5 in Sampford *et. al.*, *op. cit.*

Lambsdorff, Johann Graf. 2007. "The Methodology of the Corruption Perceptions Index 2007". Berlin: Transparency International. Available at [www.transparency.org/content/download/23965/358196](http://www.transparency.org/content/download/23965/358196) (Viewed 5 September 2008).

Leff, Nathaniel. 1964. "Economic Development through Bureaucratic Corruption." *American Behavioral Scientist*, pp. 8-14.

Manning, Nick, Dirk-Jan Kraan, and Jana Malinska (OECD GOV). 2006. "How and Why Should Government Activity Be Measured in 'Government at a Glance?'" (GOV/PGC(2006)10/ANN1).

Markovits, Andrei S. , and Mark Silverstein. 1988. *The Politics of Scandal: Power and Process in Liberal Democracies*. New York: Holmes and Meier.

MCC (Millennium Challenge Corporation). 2007. "Building Public Integrity through Positive Incentives: MCC's Role in the Fight against Corruption." Washington, DC. ; April. Available at [www.mcc.gov/documents/mcc-workingpaper-corruption.pdf](http://www.mcc.gov/documents/mcc-workingpaper-corruption.pdf) (Viewed 5 September 2008).

MCC (Millennium Challenge Corporation). 2008. *Millennium Challenge Corporation 2008 Country Scorebook*. Washington, DC: Millennium Challenge Corporation.

Miller, William L. 2006. "Perceptions, Experience and Lies: What Measures Corruption and What do Corruption Measures Measure?" Ch. 8 in Sampford *et. al.* , *op. cit.*

Moodie, Graeme C. 1980. "On Political Scandals and Corruption." *Government and Opposition* 15(2): 208-222.

*New York Times*. 2008. "States' Data Obscure How Few Finish High School." March 21. Online at [www.nytimes.com/2008/03/20/education/20graduation.html?em&ex=1206244800&en=41ef2372a3728605&ei=5087%0A](http://www.nytimes.com/2008/03/20/education/20graduation.html?em&ex=1206244800&en=41ef2372a3728605&ei=5087%0A) (Viewed 5 September 2008).

OECD (Organisation for Economic Co-operation and Development) 1996. *Ethics in the public service. Current issues and practices*. Paris: OECD.

OECD 2000. *Trust in Government: Ethics measures in OECD countries*. Paris: OECD.

OECD 2005. *Public sector integrity: A framework for assessment*. Paris: OECD.

OECD 2006. "Issues in Output Measurement in 'Government at a Glance.'" Paris: OECD.

OECD 2007. *Elements for an updated framework for integrity and corruption prevention in the public service*. OECD, GOV/PGC/ETH(2007)1.

OECD 2008. *Towards a sound integrity framework: Instruments, mechanisms, actors and conditions for implementation*. OECD, GOV/PGC/(2008)21.

Olken, Benjamin A. 2005. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." NBER Working Paper No. W11753.

Olken, Benjamin A. 2006. "Corruption Perceptions vs. Corruption Reality." NBER Working Paper 12428. Available at [www.cid.harvard.edu/bread/papers/working/126.pdf](http://www.cid.harvard.edu/bread/papers/working/126.pdf) (Viewed 5 September 2008).

Reinikka, Ritva, and Jakob Svensson. 2002. "Measuring and Understanding Corruption at the Micro Level." in Donatella della Porta and Susan Rose-Ackerman, eds., *Corrupt Exchanges: Empirical Themes in the Politics and Political Economy of Corruption*, Frankfurt: Nomos Verlag.

Reinikka, Ritva, and Jakob Svensson. 2003. "Survey Techniques to Measure and Explain Corruption." World Bank Policy Research Working Paper No. 3071. Online at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=636433](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=636433) (Viewed 5 September 2008).

Reinikka, Ritva, and Jakob Svensson. 2006. "Using Micro-Surveys to Measure and Explain Corruption."

*World Development* 34:2, pp. 359-370. Available at <http://ideas.repec.org/a/eee/wdevel/v34y2006i2p359-370.html> (Viewed 5 September 2008).

Rose-Ackerman, S. 1999. *Corruption and Government: Causes, Consequences, and Reform*. Cambridge: Cambridge University Press.

Sampford, Charles, Arthur Shacklock, Carmel Connors, and Fredrik Galtung, eds. 2006. *Measuring Corruption*. Aldershot: Ashgate.

Scott, J. C. 1972. *Comparative Political Corruption*. Englewood Cliffs, NJ: Prentice-Hall.

Shleifer, Andrei and Robert W. Vishny. 1993. "Corruption." *Quarterly Journal of Economics* 108:3, 599-617.

Shumer, S. M. 1979. "Machiavelli: Republican Politics and Its Corruption", *Political Theory* 7: 5-34.

Søreide, Tina. 2005. "Is It Right to Rank? Limitations, Implications, and Potential Improvements of Corruption Indices." Presented at the IV Global Forum on Fighting Corruption, Brasilia, Brazil, 7-10 June 2005.

Thampi, Gopakumar K. , and Sita Sehkar. 2006. "Citizen Report Cards." Ch. 12 in Sampford *et. al.* , *op. cit.*

Thomas, M. A. 2007. "What Do the Worldwide Governance Indicators Measure?" Unpublished manuscript, Johns Hopkins University (revised manuscript, January).

Transparency International. 2006. "Corruption and Governance Measurement Tools in Latin American Countries." Available at [www.undp.org/oslocentre/docs06/Corruption%20and%20Governance%20Measurement%20Tools%20in%20Latin%20American%20Countries%20-%20June.pdf](http://www.undp.org/oslocentre/docs06/Corruption%20and%20Governance%20Measurement%20Tools%20in%20Latin%20American%20Countries%20-%20June.pdf) (Viewed 5 September 2008).

Treisman, Daniel. 2000. "The Causes of Corruption: A Cross-National Study." *Journal of Public Economics* 76:3, pp. 399-457.

UNDP (United Nations Development Program) and Global Integrity. 2008. *A User's Guide to Measuring Corruption*. Oslo: UNDP Oslo Governance Center. Available at [www.lulu.com/content/2523351](http://www.lulu.com/content/2523351) (Viewed 9 September 2008).

van der Wal, Z., W.J.C. Huberts, J.H.J. van den Heuvel, and E. Kolthoff. 2006. "Central Values of Government and Business: Differences, Similarities, and Conflicts." *Public Administration Quarterly* 30:3 (Fall), pp. 314-364.

Weber, Max. 1946 ed. "Politics as a Vocation", Ch. IV (pp. 77-126) in H. H. Gerth and C. Wright Mills, *From Max Weber: Essays in Sociology*. Oxford: Oxford University Press.

Williamson, Hugh. 2004. "Hazards of Charting Corruption." *Financial Times* (20 October). Available at <http://search.ft.com/ftArticle?queryText=charting+corruption&y=0&aje=true&x=0&id=041019000362&ct=0> (Requires free registration; viewed 5 September 2008).