

# **Governance and Conflict Indicators Report**

**January 2011**

*Submitted by*



*Chris Barnett, Julian Barr, Angela Christie, Belinda Duff and Shaun Hext*

## Table of contents

<b>Acronyms and Abbreviations.....</b>	<b>4</b>
<b>SECTION 1: INTRODUCTION .....</b>	<b>5</b>
<b>SECTION 2: METHODOLOGY.....</b>	<b>7</b>
Typology.....	7
Definitions.....	7
Relevance Criteria.....	7
Data Robustness.....	11
Criteria Limitations .....	12
<b>SECTION 3: FINDINGS .....</b>	<b>18</b>
Stage 1: RELEVANCE .....	18
Stage 2: FINDINGS – country specific relevance test.....	25
Stage 3: ROBUSTNESS.....	28
<b>CONCLUDING REMARKS .....</b>	<b>31</b>
<b>RECOMMENDATIONS .....</b>	<b>34</b>
<b>REFERENCES &amp; FURTHER READING .....</b>	<b>38</b>
<b>Appendix 1: Terms of Reference.....</b>	<b>40</b>
<b>Appendix 2: Original List of Governance and Conflict Indicators .....</b>	<b>46</b>
<b>Appendix 3: REVISED Programme-Level Governance and Conflict Indicators....</b>	<b>58</b>

## Tables

Table 1: Clarity .....	8
Table 2: Ambiguous indicators.....	8
Table 3: Rule-Bound .....	9
Table 4: causally-linked.....	9
Table 5: gender or pro-poor focus .....	10
Table 6: Relevance Criteria .....	13
Table 7: Robustness Criteria .....	13
Table 8: General Relevance (Score: Strong= 1, Weak=0) .....	14
Table 9: Utility (Score - 2 = Good match, 1= similar intention, 0 = not present) .....	14
Table 10: Robustness (2=Strong, 1=Sufficient, 0 = Weak) .....	14
Table 11: Robustness Test Scale Descriptors .....	15
Table 12 (utility test) .....	16
Table 13: Scope and indicator numbers by outcome and output.....	18
Table 14: unclear measurement .....	20
Table 15: misplaced indicators.....	21
Table 16: Causally related indicators.....	22
Table 17: Cross-sector utility .....	23
Table 18a: Match between Suggested Indicators & those in the Country LFs (Outcomes) .....	27
Table 18b: Match between Suggested Indicators & those in the Country LFs (Outputs) .....	27
Table 19: Elections Indicators scoring 50% or more .....	29

## Graphs

Graph 1: Assessment of Clarity for outcome and output indicators.....	19
Graph 2: Assessment of whether indicators are Rule-driven .....	21
Graph 3: Assessment of whether indicators are Causally-linked.....	22
Graph 4: Assessment of whether indicators are Gender orientated .....	25
Graph 5: Election suite utility test.....	26

## Acronyms and Abbreviations

AFGN	Afghanistan
CAR	Capability, Accountability and Responsiveness
COR	Corruption
COS	Conflict Sensitivity
CPP	Conflict Prevention and Peacebuilding
CSO	Civil Society Organisation
CSR	Civil Service Reform
DDR	Demobilisation, Disarmament, Reintegration
DFID	Department for International Development
DOC	Addressing underlying Drivers of Conflict
DQAF	Data quality assessment frameworks
E&A	Empowerment and Accountability
ELE	Elections
GPR	Governance Portfolio Review
HRT	Human Rights
IEC	Independent Electoral Commission
IMF	International Monetary Fund
LF	Logframe
M&E	Monitoring and Evaluation
MED	Media
NAO	National Audit Office
PAS	Parliament Support
PEFA	Public Expenditure and Financial Accountability
PES	Peace Process Support
PPS	Political Party Support
S&J	Security and Justice
TAX	Tax/Revenue
UNDP	United Nations Development Programme
VFM	Value for Money

## SECTION 1: INTRODUCTION

A good results framework that enables programme progress to be effectively monitored and explained has never been more important, particularly in times of financial austerity. As a public sector organisation, DFID must have the capacity to prove that its budget is spent wisely, and the ability to demonstrate the impact and value of its programmes to core constituencies.

Governance and conflict programming form an important part of DFID's global programme portfolio and account for a significant proportion of annual resource allocation (GPR, 2010). This is likely to grow as DFID commits to expand its presence in fragile and conflict-affected environments (DFID, 2009).

It is however widely acknowledged that the effects of governance and conflict interventions on poverty reduction or enduring peace and security are seldom direct and easy to measure. International governance datasets (such as the World Governance Index), whilst comprehensive and well-resourced, seldom have relevance at actual country level as their measurements are often set at higher objective levels, yet there is a paucity of useful programme level tools available to enhance measurement in this area.

DFID began to place a greater priority on indicator selection as a means of enhancing the quality of programmatic results frameworks after the publication of its third White Paper, highlighting the importance of measuring how governance benefits the poor (DFID, 2006, DFID, 2007, DFID, 2008).

In October 2010, ITAD was commissioned by the DFID Politics and the State Team to assess the quality of a *suggested list* of governance and conflict indicators as part of a wider contract to support elements of the Results Action Plan (See TOR, Annex 1). The specific objectives for this element of the assignment are:

- *To test the relevance and robustness of the draft list of suggested indicators and to assess which are the most suitable for different programming purposes (which indicators best tell us whether we have achieved what we set out to do?). Where indicators are considered inappropriate, alternative suggestions should be provided.*

The idea of a *suggested list* emanated from the Results Action Plan, which required the production of suggested programme-level indicators for all areas of DFID activity. An initial list was drawn up by the Politics and the State Team, and it is this list that is to be tested and refined under this assignment. The list consists of fifteen separate suites of outcome and output indicators and covers a spectrum of related

programme areas, including; security & justice, elections, civil service reform, corruption, as well as other important areas where DFID allocates resources (Annex 2).

The indicators have been tested using a set of normative criteria that collectively aims to ensure the types of measurements included in the list and the corresponding data sources are fit for intended purpose. Although the study has to some extent been constrained by lack of time and available information, attention has been given to interrogating the traction of indicators with existing programme results chains and underlying theories of change, including in contexts of fragility and conflict, such as Nigeria and Afghanistan.

The report consists of three sections. The first section explains the research methodology, paying particular attention to framing a proposition to inform assessment criteria selection and a consistent scoring and transparent scoring system to test indicators for relevance and robustness. Findings are discussed in section two; empirical evidence is used to substantiate a number of observations concerning the relative strengths and weaknesses of the indicators, collectively and individually, including the identification of core indicators for each suite. Some concluding remarks are made in the final section leading into a number of recommendations for areas where additional work is needed. A revised list of suggested indicators is annexed to the report (Annex 3).

## SECTION 2: METHODOLOGY

### Typology

The suggested indicator list (Annex 2) includes fifteen separate suites identified by the Politics and the State Team; each represents a specific governance and conflict sector against which DFID currently allocates organisational resources.

Categorised as outcome or output indicators, each suite includes between 5-20 individual indicators, totalling 165 indicators. The number of outcome indicators per suite varying between 2 and 8, whilst that for outputs varies between 3 and 11. With a ratio of 5:3, output indicators outnumber outcome indicators.

Accepting that not all that can be counted counts and not all that counts can be counted, the indicator suites are squarely focused on quantifiable data, and results that can be counted.

### Definitions

A suggested indicator list is 'Relevant' if it is fit for DFID's purpose. Essentially the indicators should be consistent with, and capable of measuring, the main objectives in DFID's governance programming. In measuring objectives they will have wider utility, meeting a number of concomitant organisational priorities and commitments, specifically but not exclusively focused at country level of operations:

- Inform country and issue-based strategies
- Inform individual programme/project designs
- Inform government and policy network dialogue
- Track programme/project performance
- Forecast potential and assess actual impact
- Manage and mitigate risk
- Evidence results offers and enable business case appraisal
- Inform resource allocation decisions
- Demonstrate value for money and improving democratic accountability

Intended as part of a Results Action Plan (DFID, 2007) such a list contributes to improved aid effectiveness and helps DFID demonstrate that aid works and taxpayer money is spent wisely.

### Relevance Criteria

The list, collectively and individually, needs to exhibit a level of quality that will help country programme teams meet these different operational demands. Relevance is more about the design stage of governance projects, ensuring that they

contribute to the foundations of an effective project monitoring and evaluation system, viable indicators with life beyond project appraisals.

In a general sense, the suites should be comprised of ‘good quality indicators’. Donor guidance notes on indicator framing, whilst not always specific to governance and conflict, go some way towards identifying what this means in practice (DFID, 2009, World Bank 2004, UNDP, 2006). This initial set of criteria is described in the sections that follow:

### i) Clarity

Although donor recommendations often differ, two common principles that frequently reoccur are those of indicators that are ‘specific and measurable’ (Table 1). Good quality indicators are said to be neutral and precise units of measurement that do not set direction. Such qualities improve potential for targeting and aggregation.

**Table 1: Clarity**

Clear Indicator	Outcome	% of people who feel safe going out at night
Clear Indicator	Output	# of reported incidences of violence in the area

This type of framing advice pays particular attention to seeking clarity among individual indicators, whether at output or purpose level. How To Notes emphasise the importance of ensuring each indicator specifies what is to be measured rather than what is to be achieved. Accordingly good quality principles, indicators measure a single variable and avoid restating elements of the expected result. Terms such as ‘quality of’, ‘access to’, ‘extent of’ imply a direction and should be avoided (Table 2).

**Table 2: Ambiguous indicators**

Unclear Indicator	Outcome	Extent of respect of code of conduct by main political parties
Unclear Indicator	Output	Quality of training provided

Clarity is also a means of encouraging brief and precise indicator statements, in keeping with the requirements of organisational data management systems such as ARIES (DFID, 2009).

### ii) Rule-bound

Programme indicator framing guidance is influenced by the principles undergirding the design of logical frameworks: as objective measurements, indicators are expected



to evidence different elements of a result chain, helping practitioners explain a particular theory of change<sup>1</sup>.

As such, outcome level indicators are used at purpose level, intended to measure change among beneficiaries, whether attitudinal, behavioural or performance-based. Whilst output level indicators are measurements that contribute to demonstrating lower level results; evidencing the direct deliverables of a project. A relevant indicator list should conform to these principles (Table 3).

**Table 3: Rule-Bound**

Outcome Indicator	% citizens expressing trust in formal rule of law institutions
Output Indicator	# of cases resolved by provincial courts

### iii) Causal relationship

Accepting that the causality of good or bad governance is multiple and diffuse and the narrow linearity of such causal logics highly contestable, we cannot refute the fact that single indicators at either output or purpose level do not exist in isolation but form part of a causal relationship with a subordinate or higher level result. For this reason the output and outcome indicators included in the suggested list should illustrate such a bond (Table 4).

**Table 4: causally-linked**

Outcome Indicator	% seats in parliament held by women
Output Indicator	% women candidates (political party or independent)

### iv) Gender and Pro-Poor

Political economy perspectives caution against overly technical approaches to measuring governance change in fragile and conflict settings (Church & Rogers, 2004, Policy Practice, 2010, Tripathee, 2007). Political analysts recommend the use of indicators that measure change in the institutional structures that underpin the distribution of power and shape relations between state and societal actors. Suggested indicator suites should engage directly or indirectly with these dynamics. This also means indicators suites should measure change in those institutional

---

<sup>1</sup> A Theory of change is closely related to the implementation logic of a development intervention – the links between inputs, the implementation strategy and the intended outputs and outcomes. The Theory of Change describes the assumed or desired causal relationship between the activity or policy and its (intended) goals. Accurate and clearly stated theories of change are seen as necessary for effective programming, as well as providing a useful basis against which to monitor and evaluate performance.

structures that might empower women, the poor or other excluded societal groups (Table 5).

**Table 5: gender or pro-poor focus**

Direct outcome gender indicator	% seats in parliament held by women
Indirect output gender indicator	# criminal cases of domestic violence (disaggregated)

#### v) Political Assumptions and Risks

Theories of change, however narrowly defined in terms of explanatory variables, are still contingent on a range of stated assumptions holding true. Indicator suites that include measurements of political assumptions might also engage with the perverse incentives often generated by aid interventions and contribute to better risk management.

A good governance and conflict indicator suite will cohere with the broader conceptual frameworks that underpin programme design across different governance and conflict sectors. If this is the case, suites will contain indicators that measure capability, accountability and responsiveness as currently articulated within DFID organisational policy documentation (DFID, 2007).

#### vi) Cross-Sector Linkages

Given the conceptual convergence, it may also be expected that some suggested suites will contain indicators that resonate across sectors. For instance they may intend to measure outcomes contingent on the capabilities of a number of related institutional arenas such as political parties, elections and parliaments. They might also hope to monitor the effect of countervailing pressures between state and society such as media oversight of electoral integrity or human rights obligations. Such an integrative approach will ensure the value of the list is greater than the sum of the parts.

#### vii) Participatory

Whilst participation is an overriding principle guiding the shaping of all results frameworks, the question of who decides what to measure or in whose interests these decisions are made is important. Change is often a matter of perspective. In the context of conflict such processes are often part of the peace-building process. It stands to reason that those who have a stake in the change should be consulted on what the change looks like.

### **viii) Utility**

However, the key test of Relevance for any suite is utility: the extent of usage across those country operations with a strong governance and conflict focus. Good quality indicators suites will include a high number of indicators that closely match those already in use across these country programmes. Lack of utility might then highlight gaps in the spread of indicators currently used in the suites.

### **Data Robustness**

Relevant indicators also need to stand-up to a range of data quality assurance tests. The indicators have little value unless the data typology, data source, data collection actors and information systems are robust enough to yield evidence that is both statistically credible and defensible.

Data quality assessment frameworks (DQAF) recommend a wide range of criteria that aim to assess confidence in data sources (IMF, MEASURE, 2007; DFID, 2010). For instance, the comprehensive IMF DQAF covers five dimensions of data quality, emphasising the importance of assurances of integrity, methodological soundness, accuracy and reliability, serviceability and accessibility.

Data quality is also about how well knowledge management systems represent the real world. In many contexts information systems are weak and data can be unreliable and collection risky.

Broadly, robust indicator suites should be informed by data that is credible and trustworthy, reliable and regularly available at acceptable costs to DFID. Increasingly the data also needs to meet broader organisational priorities, for instance to be disaggregate-able (sex, age, ethnicity, geography, etc) and be aligned to wider principles underpinning aid programmes, such as government ownership.

## Criteria Limitations

There are a number of limitations that will contain the scope of the assessment criteria to what is doable given the boundaries and timeframe of the study. Some important factors include:

- The lack of supplementary information contained in the draft list makes it difficult to test for strategic coherence and to ground the suites in any underlying theories of change.
- Lack of strategic information also makes it hard to assess to what extent the indicator suites contain a balanced set of indicators that measure the three dimensions of the DFID CAR (Capability, Accountability and Responsiveness) framework.
- There is an inherent danger that indicators might be interpreted out of strategic context or from perspectives other than those intended by designers. The lack of clarification around meaning also renders the scoring process open to a certain degree of subjectivity.
- Paucity of relevant project documentation will limit the ability to assess the extent that indicator suites track political assumptions and risks.
- It is not possible to assess the views of different programme stakeholder against these types of measurement.
- The ability to test against the Robustness criteria is wholly contingent on the detail presented in the source cells of the available country logical frameworks. At this level of analysis it is impossible to assess whether, for instance, collection is institutionalised or affordable or international standards are adhered to or even if data is politically neutral.

The two tables below outline the selected criteria (Table 6, Table 7).

## The Selected Criteria

**Table 6: Relevance Criteria**

	Criteria	Definition
1	Clarity	Unambiguous unit of measurement focused on a single subject, easy to target, aggregate and communicate. Consists of three components (makes sense, neutral, precise)
2	Rule-Driven	<p>a) Outcome indicators: measure changes in attitude, behaviour or performance of beneficiaries. Evidence a shift in power relations between state and society or among different groups in society.</p> <p>b) Output indicators: measure direct deliverables, short-term results of interventions that produce better knowledge, more products or resources, greater service efficiencies etc.</p>
3	Causally-Linked	An output indicator exists as part of a results chain, in a cause-effect relationship with an outcome indicator and vice-versa.
4	Gendered	Measures structures that shape power relations between men and women.
5	Pro-poor	Measures structures that shape power relations in society, particularly factors that might empower the poor.
6	Cross-Sector	Measure has potential relevance in other governance sectors.
7	Utility	Closely matching indicators are present in a number of country programme level project logical frameworks with no significant gaps in the suites.

**Table 7: Robustness Criteria**

	Criteria	Definition
1	Available	Data source exists – primary (DFID contracted) or secondary (other organisation).
2	Accessible	Data is collectable and understandable (without high cost/unmanaged risk).
3	Credible	Data is trustworthy and defensible and does not change according to who collects.
4	Nationally owned	Data is either collected by governmental or local CSO partners, potentially as part of their own information systems.
5	Disaggregate-able	Profile information of important sub-groups can be extracted (age, gender, tribal, profession, regional).

## Assessment Test Tables

The assessment requires three tests. The first two tests for general Relevance and the third for Robustness. The first test assesses all suites of indicators against six of the seven Relevance criteria (Table 8), the second tests for Utility (Table 9), and the third test is against all five Robustness criteria (Table 10). Each table may be adopted to test either outcome indicators or output indicators.

**Table 8: General Relevance (Score: Strong= 1, Weak=0)**

Indicators	Clarity	Rule-Driven	Causally-Linked	Total	Gender	Pro-Poor	Cross-Sector	Total
Total								

**Table 9: Utility (Score - 2 = Good match, 1= similar intention, 0 = not present)**

Indicators	C 1	C 2	C 3	C 4	C 5	Total
Total						

**Table 10: Robustness (2=Strong, 1=Sufficient, 0 = Weak)**

Indicators	Available	Accessible	Credible	Own-able	Disaggregate-able	Cross-Sector	Total
Total							

## Scoring System

A binary scoring system is used for the first test (Table 8) and a three-point scale for the second and third (Tables 9 & 10). The binary score used in the first Relevance test allocates 1 for presence of an indicator as defined and 0 for indicators that do not conform to the criteria descriptions.

The Utility test three-point scale ranges from 0 for an indicator that is not used in the country programmes, 1 for an indicator that may be worded differently but intends to measure a similar result, and 0 for an indicator that is not present in any of the tested country project LFs.

The Robustness assessment scale requires the indicator to be present in one or more selected cluster LFs and is predicated on the assumption that some indicators will have precise information presented in the corresponding source cell of the country LF, a sign that country teams have a clear idea of what data source and data type will be used to populate the indicator. These will be considered strong and scored 2

whilst those with only partial information will score 1 or 0. A scale descriptor is used to enhance consistency in this process (Table 11).

**Table 11: Robustness Test Scale Descriptors**

Scale Descriptors			
	Strong (2)	Sufficient (1)	Weak (0)
<b>Available</b>	Data source named. Source cell may be split into two, clarifying collector	Several data sources but unclear which are most important or produces or collects	Insufficient detail
<b>Accessible</b>	DFID or like-minded actor is main producer and collection agency Data is in recognised format	Does not specify who collects data and unclear if format will be easily understandable	Insufficient detail
<b>Credible</b>	Data is from a trusted international known for best practice, methodological shortcomings are known	Not sure on credibility of source, probably unknown shortcomings.	Insufficient detail
<b>Own-able</b>	Data type is embedded in partner government or CSO information systems.	Joint donor-government data source, contingent on external funding and extension of project	Data type and source is specifically established as part of donor M&E system with few linkages to government or CSO networks
<b>Disaggregate-able</b>	Data type is clearly disaggregate-able by more than one sub-group, which are specified in source cell	Disaggregation is implicit but limited sub-groups such as gender	Data typology does not appear conducive to any form of disaggregation

The robustness test also serves to substantiate findings from the utility test, highlighting those indicators that are not in use at the project level (Table 12).

**Table 12 (utility test)**

Governance Arena: Elections																		
Criteria		Available			Accessible			Credible			Own-able			Disaggregate-able			Total	
Scale		S	A	W	S	A	W	S	A	W	S	A	W	S	A	W		
		2	1	0														
Output Indicators																		
#1 = legal framework in place (up to international standards)																		
AFGN: election regulations and guidelines formulated by IEC	Key and	2			2			2						1			0	7

All criteria have been accorded equal value with no use of weightings. There are however some implicit differences. Three Relevance criteria (clarity, rule-bound, causally-linked) are considered technically essential qualities: good practice that should be followed in the framing of all governance and conflict indicators. Whilst three other relevance criteria (gender, pro-poor and cross-sector), should be considered as desired qualities *of the suites* and the suggested set as a whole. These latter qualities are important as governance programmes enable gender and pro-poor outcomes rather than specifically target them - an important distinction explored in the findings section.

**Assessment Limitations**

The selection of assessment criteria is limited to what is achievable within the study timeframe. For instance, it is not possible to explore the extent that indicators cohere with underlying theories of change, effectively integrate political drivers, track perverse incentives or manage project risks, nor is it possible to assess the scope of participation.

Whilst the first test for Relevance can be undertaken for all 15 indicator suites, the Utility test is a process of assessing to what extent suggested indicators are actually used in country project LFs. The test is contingent on a reasonable sample of LFs being available from a given sector. Whilst there is no shortage of LFs, there is however a dearth of LFs that cohere with a given sector. For this reason the second and third tests have been limited to three sectors where information is available: Tax and Revenue, Elections, and Security & Justice. Such findings may not be broadly applicable.

As some suites are limited to two or three indicators, caution should also be exercised when comparing percentage differences between sectors. Moreover



absence of PEFA indicators in the assessment may also distort findings from some suites that include PEFA indicators.

Despite the use of descriptors, some educated guesses are called for. For instance at this level of analysis we cannot know for sure that data is disaggregate-able or whether data is government owned. Interpretation of Robustness findings will need to be exercised with caution.

## SECTION 3: FINDINGS

A three-stage assessment process was used. The first stage tested the 15 suites and all 166 outcome and output indicators (with the exception of six PEFA indicators) against five general relevance criteria. Due to limited time and LF availability it was agreed that the second stage (Utility) and the third stage (Robustness) tests would only include three suites: Tax and Revenue, Elections and Security & Justice (Table 13). Findings from each stage are presented below. Select graphs and tables are used to highlight the overall messages only. Examples of corresponding good or bad indicators are included to substantiate the main points. A general health warning precedes each sub-section.

### Stage 1: RELEVANCE

#### Health warning

The figures need to be interpreted with a degree of caution as it can be misleading to use percentages to compare suites with significant spreads in the numbers of indicators (Table 13). Moreover there is always a risk that percentages will distort the real picture when numbers are very low, such as those for Media, Human Rights. That said, although the overall list lacks a certain balance in terms of indicator numbers randomness is to be expected from draft suites that have so far benefitted from varying amounts of design input.

**Table 13: Scope and indicator numbers by outcome and output**

	S&J	CSR	ELE	PAS	PPS	MED	TAX	COR	HRT	PPS	DDR	CPP	DOC	COS	E&A
OC	6	6	8	6	5	2	3	5	3	3	3	3	6	2	2
OP	11	11	12	5	6	7	5	5	6	7	6	8	4	3	7

#### Overview

Broadly the suggested list exhibits distinct pockets of quality. Although there is certainly an imbalance: several suites contain outcome and output indicators that are noticeably stronger against all six relevance criteria than others. For instance outcome indicators for Security and Justice (69%), Tax (67%), and Elections (64%) generally outperform outcome indicators for Demobilisation, Disarmament, Reintegration (38%), Political Party Support (34%), Conflict Sensitivity (29%) and Human Rights (29%).

These general findings support more widely held views that governance and development sectors that are easier to measure tend to perform better than those more difficult to pin down (Natsios, 2010). Accordingly, sectors that contain more measures will often perform better than those that contain fewer; those

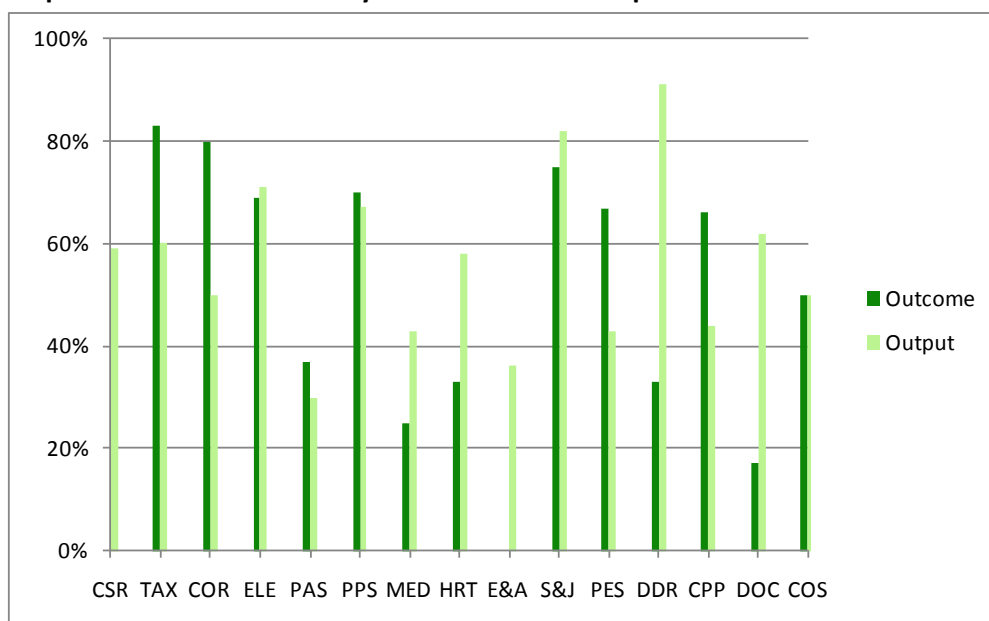
institutionally bounded will be easier to measure than those more institutionally diffuse; and technical measures are usually easier than to track than political ones.

### Why can clarity be so hard to pin down at outcome level?

A common dimension of indicator quality is clarity. Unambiguous specific units of measurements are important as they enable programme effectiveness needs to be analysed, valued and communicated to both external and internal audiences (NAO, 2010, DFID, 2010).

Many of the suites contain indicators that conform to our definition of clarity (Table 6). Those sectors that are easier to measure: Tax (83%), Corruption (80%), Elections (69%), Political Party Support (70%), and Security & Justice (75%), tend to contain clearer measures at outcome level than those listed in more difficult to measure areas such as Drivers of Change, Conflict Sensitivity, Media, Empowerment and Accountability. This may be research distortion, a result of aggregating percentages against low sample sizes, but it could also be indicative of a more general problem governance project designers have with nailing clear outcome measures in more diffuse sectors, particularly as related output indicators score significantly higher in some of these sectors (Graph 1).

**Graph 1: Assessment of Clarity for outcome and output indicators**



Problems with outcome clarity is often be attributed to over-ambitious programming, coupled with the desire to comply with organisational advice that cautions against the use of too many indicators against each objective (DFID, 2009). Whilst designers seem to experience less problems at output level, at outcome level they tend to package related aspects of change as multi-variant indicators, worried

that information will get lost if they opt for too specific a unit of measurement (Table 14).

Such behaviour might be viewed as a form of resistance to the techno-centric and narrow linearity of results frameworks, an understandable attempt to carve out some wriggle room, a refusal to accept that complex change can or should be evidenced by a single unit of measurement.

**Table 14: unclear measurement**

Unclear outcome indicator	Average nos. in cells, prevalence of infectious disease/malnutrition; nos. Rapes per capita prison population
Clearer outcome indicator	Number of reported infectious diseases among prison population

This behaviour is also found in a similar tendency to conflate indicators with objectives. Ambiguous terms such as ‘level of’, or ‘access to’, ‘existence of’, ‘mechanisms in place’ are evident across seven of the fifteen suites (S&J, PPS, PAS, HRT, DDR, DOC, COS). In cases where governance project designers, often with high level ambitions, find themselves restating their objectives, it is often a sign that their objectives require greater specification. Sometimes the practice might be construed as a sign that governance and conflict assessment findings have not been sufficiently factored into the design process.

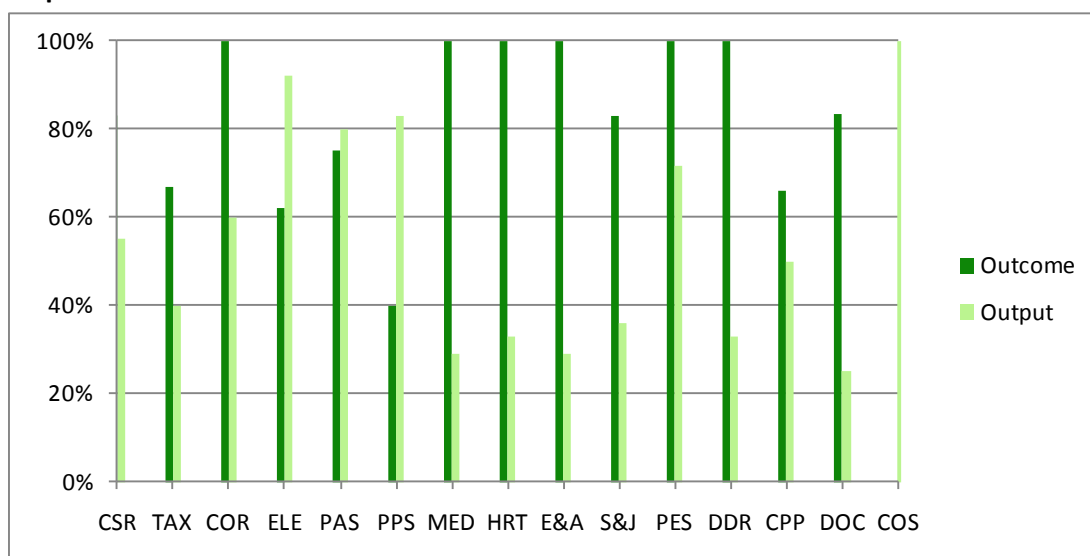
For instance, to use an example from the DOC suite (Annex 2): “% minority access to government employment”. The reason for the measurement is clear, but what result is the indicator intending to evidence, implicit to the measurement? Is it that government is inclusive, less chauvinist, ethnically or tribally biased? If the issue is inclusion might it be more specific to measure the “% of minorities employed in higher level civil service grades” (assuming a supportive political environment for data collection).

#### **Why is it so difficult to differentiate between outcomes and outputs?**

As different elements of a results chain, governance indicators, whether at output or outcome level, are always context specific, highly interpretive and hence contestable (Carothers, 1999, Williams A & Siddique, A 2008). Contestation is naturally increased when indicators are atomised, disconnected from objectives or not coherently linked in a cause-effect or means-end relationship.

Despite these understandable design constraints, all suggested suites contain indicators that generally perform well against the working definition of outcome and output indicators (Table 6). There are a number of incidences where outcomes would be better placed as outputs and where outputs would be better placed as outcomes though there are no discernable patterns (Graph 2).

**Graph 2: Assessment of whether indicators are Rule-driven**



The findings suggest that perception surveys, measurements of higher level attitudinal change, are at times being presented as output measurements (Table 15). Such placements generally go against the standard definition. For instance from the DDR suite (Annex 2), the indicator “% of community who feel threatened by the presence of ex-combatants” is likely to yield responses influenced by a range of factors beyond the direct control of the project. Equally, from the Elections suite (Annex 2), a project might be directly accountable to the “# of judges trained in dispute resolution”. Both indicators are better placed at different levels.

Are there reasons why such outcome indicators often find themselves at output level? It may for instance be a characteristic of governance and conflict programme design, a result of over-ambitious plans in a complex programming environment. Under these conditions change is often pushed down to output level. An example of this dynamic can be found in the use of cascade approaches, such as those used by international agencies to nest multi-component logical frameworks. Such processes often produce unintended consequences as one organisation’s outputs, with no intended direct transformational effect, become elevated to another’s outcomes or purpose with greater accountability for change.

**Table 15: misplaced indicators**

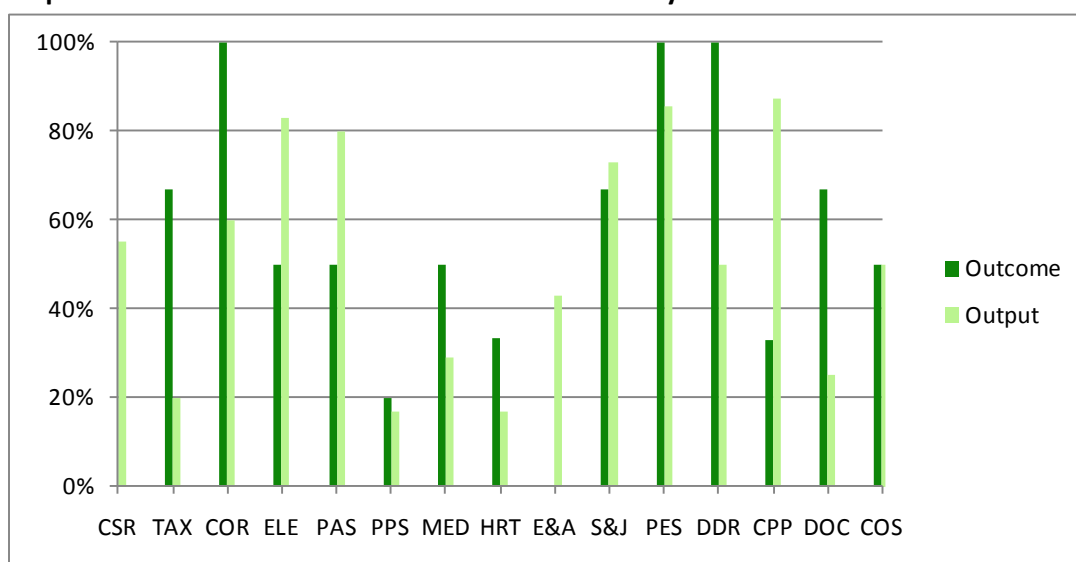
Misplaced Outcome indicator	# of judges trained in dispute resolution (Elections)
Misplaced output indicator	% community who feel threatened by presence of ex-combatants (DDR)

### Why should outputs and outcomes be causally-linked?

An underlying assumption of this study is that coherent indicator suites, those whose output and outcome indicators are causally related, is likely to have greater value than one that is little more than the sum of its various parts. An integrated suite of indicators is considered better than one where the indicators are somewhat disconnected from any direct causal logic (Table 16).

There is certainly evidence that several suites contain output indicators that are causally related to corresponding outcomes, for instance Conflict Prevention and Peace Building (87%), Peace Process Support (86%), Elections (83%), Parliamentary Support (80%) and Security & Justice (73%). Even lower scoring suites such as Political Party Support (17%) contain some output indicators that exist in a credible means-ends relationship with a related outcome (Graph 3).

**Graph 3: Assessment of whether indicators are Causally-linked**



**Table 16: Causally related indicators**

Outcome indicator	% political party accounts publicly available
Output indicator	% political parties producing annual plans and budgets

However it is difficult to ascertain whether this arrangement is by design or by accident. The general lack of traction between the two levels of a hypothetical results chain suggests that suite designers have paid more attention to crafting individual indicators rather than weaving indicator baskets.

To be sure, it is important for indicators to be specific and measurable, even for them to be participatory and empowering, but it is also critical that they do not exist in isolation.

In much the same way that it can be dangerous to design interventions without input from others that have a stake in the outcomes. It is also critical that the indicators collaborate, that they are part of a bundle that collectively paints a rich picture of how change happens and what change looks like. This is ever more important if the drivers of change are understood in a less political sense, framed as part of a linear and planned process rather than the product of more complex dynamics.

### Why are linkages between suites weak?

Governance and conflict programme sector measurements are not distinct bounded entities. In some instances they are linked through institutional ties such as the natural relationship that exists between support for political parties, elections and parliaments. In these cases indicators such as those measuring electoral accountability, integrity or inclusion will have cross-sector utility.

In other instances indicators may collectively contribute to measuring cross-suite governance outcomes. For instance, in evidencing the accountability function of CSO's and media organisations, capturing rich descriptions of what countervailing pressures on state institutions look like (Table 17).

One might also expect to spot evidence of inter-sector causal linkages around engagement with structural drivers. For example, in the way that a DDR output measurement (Annex 2), “% ex-combatants able to maintain an independent livelihood”, has potential links to a DOC outcome measurement (Annex 2) – “% unemployment of youth”; framed as they are around recognition of the same economic drivers of conflict.

**Table 17: Cross-sector utility**

Media/Elections	# reports of electoral integrity in national media
S&J/E&A	# CSOs consulted on policy development

Whilst the findings evidence the potential for traction (particularly for security and justice, conflict prevention and peacebuilding, elections, empowerment and accountability, and drivers of change suites), it is unrealistic to expect this sort of value added to exist in early drafts of the indicator sets. It may be reasonable to expect to pick-out evidence of inter-suite links without a conscious design input, perhaps as a natural outcome of the similar causal logics that pervade the suites, but

it is also possible that such traction might be impeded by design, particularly if there are evident imbalances in terms of causal measurements.

To illustrate: too many capability measures might produce indicators with a technical rather than relational focus. Whilst an overemphasis on capability might reflect an interest in institutional performance it may also lead to accountability measures that concentrate on a specific state or societal institution rather than an accountability relationship between both institutions.

#### **Why must gender and pro-poor indicators amount to more than disaggregation?**

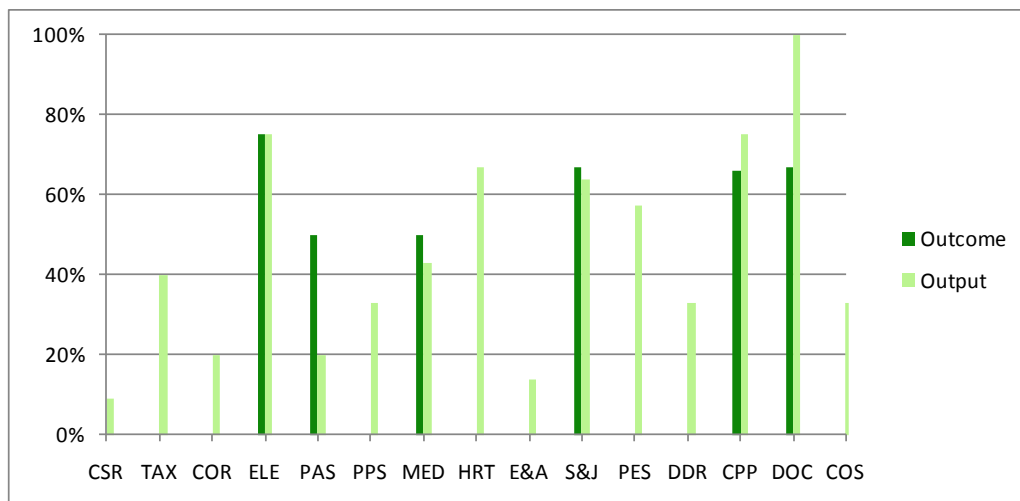
As an organisational imperative governance and conflict indicator suites should measure changes in structures that shape power relations in society, specifically those that empower the poor or drive the often uneven distribution of power between men and women or ethnic groups (DFID, 2007).

Bearing in mind the earlier caveat regarding small samples sizes, findings suggest that whilst there is some evidence of indicators with gender or pro-poor dimensions the focus is generally sparse at higher level (Graph 4). In fact suites such as tax, corruption, human rights and political party support contain no outcome indicators that could be interpreted as gender or pro-poor (Annex 2).

The finding might also have its roots in organisational practice: particularly if gender and pro-poor measurements are perceived as a bureaucratic process of data disaggregation. Of course, disaggregation is not without value, data profiling can serve a gender specific or pro-poor purpose, particularly when assessing the inclusiveness of state institutions, for example “% women in parliament or % minority candidates”. However the use of disaggregation can also cause indicators to be narrowly interpreted, for instance from DDR (Annex 2), “# of combatants disarmed and demobilised” might be disaggregated to identify the ethnic dynamics of DDR processes but in so doing overlook its proxy nature in terms of gender-based violence, suggesting a broader approach may be required (UNDP, 2006).



**Graph 4: Assessment of whether indicators are Gender orientated**



The practice of tracking gender and pro-poor results via disaggregation may also have its roots in what has been recognised as a tendency for DFID governance programmes to overemphasise interventions that aim to enhance state capability and institutional performance (DFID, 2010).

Measures that illustrate change in the structural factors that shape societal relations are less likely to be framed if the entry point is capability. Conversely measures of accountability, whilst more aligned to such structural changes, for instance through the introduction of legislative reforms that might empower women, are more political and hence difficult to track and deliver.

Suites would likely be more responsive to measuring gender and pro-poor change if took a broader approach (UNDP, 2006).

### **Stage 2: FINDINGS – country specific relevance test**

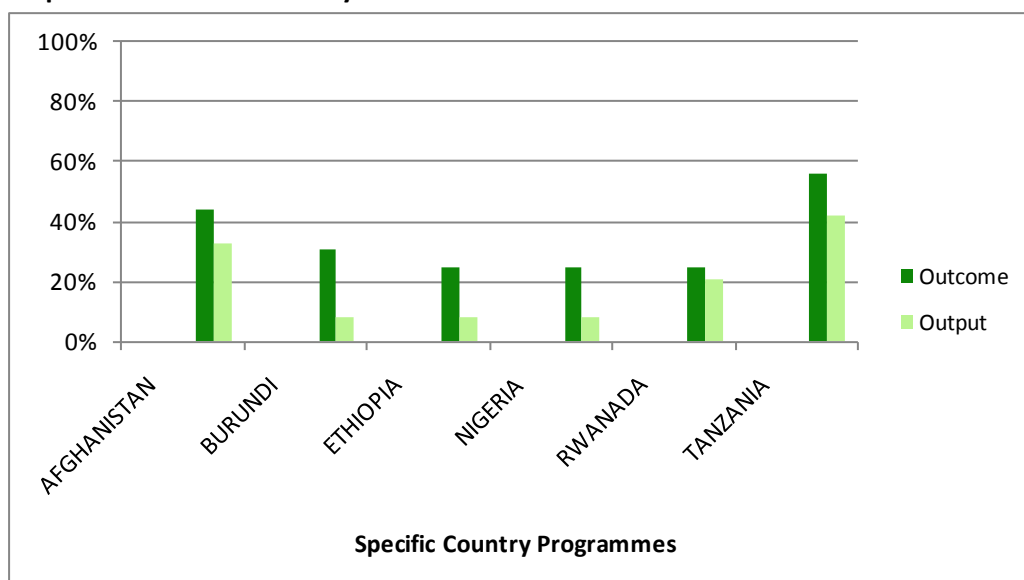
The second stage test for Relevance is a test for utility. It is essentially a matching process, to assess traction between three suites (tax and revenue, elections, and security & justice), and output and outcome indicators used in thematically related project LFs in use at country programme level.

The election suites were assessed for utility against six country LFs: Afghanistan, Burundi, Ethiopia, Nigeria, Rwanda and Tanzania. Tests for the Tax and Revenue suite drew on logical frameworks from Afghanistan, Bangladesh, Pakistan and Sierra Leone. Tests for Security & Justice compared the suite indicators with projects from Bangladesh, Burundi, DRC, Malawi, Nigeria and Southern Sudan.

### Why did some parts of the suite have greater traction than others?

The Election indicator suite generally exhibited a good match with the six country LFs tested, particularly in terms of outcome indicators. Those indicators with the highest number of matches were with countries with a discrete focus on elections: Afghanistan and Tanzania (Graph 5).

**Graph 5: Election suite utility test**



The reason for the lower score against other countries can be attributed to one key factor. The test was not undertaken on a like-for-like basis. In many countries, governance projects are often structured as wider democracy-building projects, engaging in partnerships beyond electoral commissions and political parties and issues beyond electoral capability and legitimacy.

That said, several indicators that were used in country project LFs could not be matched to the suite and their presence in more than one country LF points to gaps in the Elections indicator suite. These are related to electoral participation and accountability rather than capability and legitimacy.

Outcome indicators performed better than output indicators, perhaps because the spread of results at this level is narrower and hence less room for measurement divergence. However, both outcome and output indicators recorded some close matches and those that achieved three or more can be considered important (Tables 18a and Table 18b). The reason that some of the indicators did not score is sometimes because they are misplaced, as either outcomes or outputs.

**Table 18a: Match between Suggested Indicators & those in the Country LFs (Outcomes)**

	SUGGESTED OUTCOME INDICATOR	COUNTRY RELEVANCE SET RAG RATING
1	% external observation reports stating elections have been conducted freely and fairly	4
2	% voting age registered to vote	0
3	% registered voters who vote	3
4	% satisfaction with conduct of elections (disaggregated)	3
5	% citizens received voter education (on process, rights & responsibilities)	0
6	% citizens who feel able to cast their vote without pressure	0
7	% citizens expressing confidence in capacity of police to prevent electoral violence (disaggregated)	3
8	Number of judges trained in electoral dispute resolution	0

**Table 18b: Match between Suggested Indicators & those in the Country LFs (Outputs)**

	SUGGESTED OUTPUT INDICATOR	COUNTRY RELEVANCE SET RAG RATING
1	Legal framework in place (up to international standards)	3
2	% electoral body staff trained in their specific role with job descriptions (disaggregated)	8
3	% women/minority members represented in parliament	0
4	% women/minority candidates	2
5	% population aware of election (disaggregated)	0
6	% population aware how to vote (disaggregated)	7
7	% observers trained (disaggregated)	1
8	% elections preparations completed on schedule	2
9	Number of incidents of political violence reported in national media (pre-election period, election day, post-election)	0
10	% electoral appeals concluded	3
11	Ratio national: international observers	2
12	Ratio observers: population size	0

### Why is a diverse indicator typology important?

The country-level indicator typology includes fact, perception and opinion-based measures with the latter usually more prominent at higher outcome and goal levels. The suggested indicator list places greater importance on perception-based indicators, particularly at output level than do country level LFs.

Perception indicators are important, particularly when assessing changes in citizen viewpoints but they often need to be balanced with fact-based proxy indicators to gain substance.

At the country level there also seems to be a creeping tendency to use indicators from international data-sets at purpose level. Both behaviours suggest that country indicator designers often require more support to help them craft a more diverse typology of indicators at purpose level.

### Can all that counts be counted?

A paucity of available logical frameworks from other sectors meant that Tax and Revenue, Elections and Security & Justice suites were all self-selecting. It is perhaps telling that the same three indicator suites were among the largest number of indicators and achieved the highest first stage relevance test scores. It would not be unreasonable to predict that they would likely exhibit the tightest traction with country project LFs were it possible to test all suites for Utility.

In the knowledge that not all governance results can be counted, will increased requirements for better results measurement have a narrowing effect on DFID programme portfolio? In this climate of accountability how to ensure that the operational incentives are configured in such a way that hard to measure yet important projects still happen. Do hard to measure areas need special attention?

## Stage 3: ROBUSTNESS

### General health warning

These findings should be read with caution and not taken in isolation of the previous findings. Specifically, criteria scores against data availability, accessibility, and credibility are all highly correlated and the limited depth of analysis inhibits the level of granularity necessary to distinguish the three qualities of robustness. An inability to probe the quality of the source information also undermines the objectivity of scoring against government ownership and disaggregation criteria.

### Overview

That said, The analysis can draw on a number of similarly framed indicators in use at country programme level and the source cell information corresponding to those indicators is deemed relatively robust. Specifically four of the eight outcome indicators and six of the twelve output indicators for Elections were considered robust, with averages of 50% or more across the five criteria tested (Table 19). However absence of suggested indicators across six LFs meant that two suggested outcome indicators (# 2 &7) and six suggested output indicators (#3, 5, 7, 9, 10, 12) could not be tested.

**Table 19: Elections Indicators scoring 50% or more**

Elections		
Robust Outcome Indicators	1	% election observation reports stating that elections have been conducted freely and fairly
	3	% registers voters who vote
	4	% satisfaction with conduct of elections
	5	% citizens received voter education (disagg)
Robust Output Indicators	1	Legal framework in place
	2	% electoral body trained in their specific role
	4	% women/minority candidates
	6	% population aware how to vote
	8	% elections preparations completed on schedule
	11	Ratio national: international observers

For S&J for example, only one country (Nigeria) and one indicator was considered robust at both outcome and output level. It is difficult to draw wider implications across the suite, but since Elections, Security & Justice and Tax & Revenue are already recognised as stronger suites we might suggest that harder to measure suites are likely to show less traction at country level and lower levels of robustness.

#### Concerns about sourcing robust outcome data

The cell information implies that much of the output indicator data can be sourced via DFID project management systems whilst outcome data is often contingent on external secondary sources.

Donors often consider primary sources as more reliable than secondary sources, particularly in fragile and conflict-affected environments. This may be one reason why some country teams prefer to use expert opinion data from recognised international sources such as World Bank, Freedom House, Polity IV or Transparency International at purpose level, despite its more general use at impact level.

Difficulties sourcing outcome data may also be a reason why secondary sourced perceptions surveys are becoming more popular, particularly those used by Afrobarometer, Asia foundation and increasingly issue-specific longitudinal studies commissioned by donor groups, particularly in conflict environments. Data quality in these cases is best interrogated on a case-by-case basis, often limited to qualitative assessments undertaken by programme statisticians or results advisors with a thorough knowledge of potential survey design flaws and corresponding mitigation strategies.

### **Preferences for more reliable international perspectives and data sources**

There is always a risk that the primacy afforded to international expert opinion, even if via on-the-ground electoral observer reports, might de-emphasise the importance of also measuring the attitudes and actions of important government and societal actors, in the case of elections, such as electoral commissions, political parties, electoral candidates and the media.

These risks are to some degree substantiated by the fact that only 20% of tested outcome data sources engage with local information systems. In this case, the demand for reliable data may act as a perverse incentive, on the one hand causing donor staff to establish parallel data collection systems to measure change but on the other only measuring that change where data is readily available and easy to collect, a feature of other bi-lateral programmes (Natsios, 2010).

In the context of elections, there would be a real benefit in thinking more broadly about what free and fair elections means to others with a stake in the process. A more nuanced approach would help identify additional locally sourced proxy measures that would substantiate international findings.

### **The need go beyond gender and pro-poor 'disaggregation'**

Disaggregation tests are statistically unreliable as not all measures are intended for disaggregation nor would they benefit from disaggregation; and, there is clearly more of a value in disaggregating some indicators than others. At this level of analysis it is difficult to truly know whether many of the indicator data sources can measure the effect of project interventions or wider structural change outcomes on excluded groups, such a women or the poor, in any given country contexts.

Judgements regarding the quality of the data source suggest that the potential for disaggregation is average (42%) for Election outcome indicator suite, and weak (21%) for the output indicator suite. These findings contain two important lessons. First, that the technical act of bracketing disaggregation in an indicator cell of an LF in no way guarantees the data eventually sourced will be disaggregate-able. Second, in terms of gender and pro-poor data there is a real need to go beyond disaggregation and use mixed typology of indicator including proxy measures (UNDP, 2006).

## CONCLUDING REMARKS

The study identified strengths and weaknesses of the suggested indicators list and noted areas where individual and collective indicator quality might be enhanced. The study has also produced findings that have utility beyond indicator refinement and might help inform the evolving results agenda and help to ensure that results frames continue to provide evidence to measure impact and value.

In the section that follows, we draw together the main conclusions of the study:

1. **It is essential that the selection of indicators is grounded in a clear theory of change** – especially so that they can tell a convincing story. For example, the evidence presented from Elections shows that suggested indicators are more likely to measure electoral *capability*, than accountability and responsiveness. Indeed there is a tendency across the list to focus on a narrower, or essentially normative, theory of change – one that is primarily concerned with building capability aspects of governance rather than taking account of wider dimensions of the CAR framework, or beyond it. In this way, the list conceptualises governance as primarily about a technical democracy-building process, and understates those measurements that seek to engage with political drivers or understand how more informal institutional relations shape governance outcomes. There is a tendency to use indicators to sketch artificial boundaries when the realities are often blurred, not least in the relationship between governance, growth and security.
2. **What counts is not necessarily easy to count.** The list goes some way towards highlighting what is countable but at times fails to highlight what counts. The increasing demand for reliable data may produce unintended consequences such as by causing donor/ programme staff to establish new data collection systems, which may run in parallel to existing M&E systems. Alternatively, the increasing demand for reliable data may lead to only measuring change where data is readily available and easy to collect.
3. **Attention needs to be paid to weaving a balanced suite of indicators.** In reviewing the list and DFID programme logical frameworks, it is clear that more attention has been paid towards crafting individual indicators rather than weaving ‘indicator baskets’. It is also critical that indicators do not exist in isolation, and this is ever more important when drivers of change are understood in a less political sense - framed as part of a linear and planned process rather than the product of more complex dynamics. There are many aspects to balancing a suite of indicators, but two important aspects are:

- Citizen-based perception surveys and expert opinions are both important ways of measuring governance outcomes but their credibility is enhanced if balanced by fact-based, proxy measures. The balance is not necessarily evident in the suggested list or among tested country logical frameworks.
  - There may be a tendency for donors to consider primary sources as more reliable than secondary sources, particularly in fragile and conflict environments. In reviewing the logical frameworks, this may be one reason why some country teams appear to prefer expert opinion data from recognised international sources (such as World Bank, Freedom House, Polity IV, or Transparency International) at purpose level. This is despite the more general applicability of such sources at the goal/ impact level.
4. There is a real need to go beyond disaggregation and use a mixed typology of indicators, including proxy measures to assess shifts in power relations. **Data disaggregation is important but it is not the sole means of measuring shifts in unequal social relations.** The current practice of using disaggregation as the primary lens through which to measure gender and pro-poor outcomes may also be a product of a narrow focus on state capability. If the entry point for a governance intervention is capability, gender measures are more likely to be institutionally bounded using profile data to assess institutional change. Greater focus on accountability and responsiveness will create more appropriate entry points to assess structural changes that affect power relations. For instance through measuring legislative reforms that might empower women or decision making by women.
  5. **The purpose, use and validity of indicators remains contested. There are a number of practical lessons that can be inferred from the findings.** Firstly, outcome indicators often find themselves at the output level - a feature of governance and conflict programme design, and a result of over-ambitious plans in a complex programming environment, and one where attribution for behavioural change can be difficult to establish. Under these conditions change is often pushed down to output level and the ability to track a theory of change from intervention to consequence is compromised. Secondly, output indicators are also found at the outcome level, suggesting some confusion/inconsistency over definitions and the hierarchy of the results chain. And lastly, programme designers (or those completing logframes, such as managing agents or NGOs) don't have a similar understanding of why and how to use indicators. In practice designers create measurement ambiguity rather than clarity and opt for more rather than few, with no real intent that data will be gathered accordingly. A serious gap persists between the use of



indicators for strategic planning purposes and associated securing of resources, and their wider utility as part of programmatic results frames and learning processes.

## RECOMMENDATIONS

The revised indicator list (Annex 3) provides a modification of the suggested list based on the review of the relevance and robustness of the indicator set. This has resulted in some deletions and modifications (based on the Stage 1 assessment), with additional indicators based on the review of the country LFs (Stage 2 and 3 assessments).<sup>2</sup> We have deliberately limited the further expansion of the list to those indicators that derive from the review process. Nonetheless, **we can see areas where the list could evolve further**, as shown in Box 1, and in consultation with Governance and Conflict Advisors, this might be a useful area for further work.

### **Box 1. Examples of expanding the revised indicator list**

For Human Rights for instance, instead of the original five purpose/ outcome indicators, the list might also include ones such as:

- % UPR civil and political rights recommendations implemented
- % citizens believe HR body is improving human rights situation (disagg.)
- % citizens who believe they are free to express their own religious beliefs
- State prosecutions for domestic or sexual violence as % of all reported cases
- % reported cases of political/ sexual violence investigated by government
- % citizens who believe CSOs and media protect their human rights

Additional output level indicators might include:

- # CSO and media involved in tracking UPR processes
- # parliamentary sub-committees with responsibility for oversight of human rights protocols and security reform

Plus, apart from the revised indicator list, **the study has highlighted the following additional areas** which might support improvements in the governance results agenda:

1. **A research study that explores whether there is sufficient commonality to devise a number of 'model theories of change' across the Governance portfolio.** Governance is acknowledged as a hard to measure sector. It is little understood whether it is even possible to derive a number of model theories of change that apply at least in some governance sub-sectors and which could be usefully adapted and refined to different operational contexts. Such a research study would provide the Governance cadre with an informed basis to decide about some of the risks and benefits of moving towards a more harmonised approach to results measurement.

---

<sup>2</sup> Additional indicators are provided for the three sectors of Elections, Tax & Revenue, and Security & Justice.

2. **Governance and conflict intervention designers need guidance that succinctly shows them how to move beyond the specifics of crafting individual indicators.** This should support a shift towards the use of baskets (or suites) of indicators, where a balanced set of multi-dimensional indicators are used to measure performance. For this, we are not advocating in favour of distilling change down to a narrow set of technical variables, nor the crafting of econometric baskets that all total 100%, but merely looking to weave rich information streams that suitably help understand the complexity of change. Such streams imply a mixed typology of technical and political indicators at both output and outcome level, that collectively draw on facts, opinions, perceptions and assumptions to capture knowledge, attitudes and actions among a diverse set of political actors in a given set of often overlapping programmatic contexts. A diverse suite of complementary indicators will have a number of advantages, including: (i) Firstly, that a balanced set supports validation/ triangulation, such as when combining perception-based indicators with more fact-based measures. (ii) And secondly, a combination of indicators should provide more than the sum of each individual part – and in doing so contribute to telling the real ‘story’ of change.
3. **There is a need to develop other types of indicators to unpack power relations (gender, pro-poor, etc).** The current suggested indicator list does not have enough indicators to unpack critical changes in power/ societal relations. Within the current list gender is mostly considered in terms of disaggregation rather than shifts in power relations – with the former approach tending towards *all* indicators needing to have a disaggregated ‘gender’ element, with the latter approach emphasizing that a ‘basket of indicators’ should have *some* indicators that measure the power shifts in gender relations. Possible ways to complement the suggested list, include:
  - Reviewing the current portfolio of logical frameworks and capturing how they currently measure shifts in power relations. Based on this analysis, it should be possible to consolidate and develop appropriate measures. This is a task that should probably be done jointly between Governance and Social Development Advisors.
  - Applying a gender/power analytical framework to further analyse the list of suggested indicators and test against country logical frameworks. The aim would be to identify a mixed typology of direct and indirect (proxy) indicators across core governance sectors – with ones that are more suitable to measure the effect on women/ men and other excluded groups.
4. **There should be more guidance to support the improvement in the robustness of datasets.** Many of the logical frameworks reviewed seem to

have been written with little attention to utilise the data even where the data source is quoted. There are a number of areas where data quality could be improved, with better guidance. These include:

- **Produce a 'How To Note' to advise on methodologies to deliver cost-efficient, credible and defensible indicator measurement data in fragile and conflict environments.** The quality of governance and conflict results frameworks is highly contingent on baseline data. There should be a particular focus on how to produce baseline data for fact and perception -based indicators. The emphasis would be on a process of largely synthesising what is already out there, drawing on the DFID statistician cadre networks and learning.
  - **Include a checklist to help project teams decide about the data robustness of the indicators they select.** Use consultation process to undertake deeper assessment of data robustness drawing on a wider country sample and looking at documentation beyond logical framework source cells.
  - **Produce a guidance paper on the value of different international datasets,** their shortcomings and key governance indicators that have utility in terms of programme performance.
5. **There is a need for the suggested list to expand beyond measures of capability to include accountability and responsiveness,** in a political as well as technical sense. There is a tendency to labour on perception-based indicators (“% satisfied with...”), particularly to capture the views of ‘citizens’ or ‘populations’. A commensurate absence of fact-based data, often proxy measures empirically measuring change in government actions suggested more work is still required to achieve this balance.
  6. **There is more work to be done to model and guide the horizontal logics against key governance and conflict sectors, particularly at outcome level.** In the suggested indicator list, some indicators are better placed as milestones and there is some confusion around framing indicators, particularly ensuring neutrality and avoiding setting direction. Programme designers will greatly enhance results frameworks if they can create rich information streams that illustrate qualitative distinctions between indicators, milestones and targets.
  7. **Further work is needed to better understand the extent to which governance elements of sector programmes are captured (or not) within DFID’s results system.** It may be that the governance agenda is under-reported (and

insufficiently measured) when it is part of a sector programme such as those on education, infrastructure and health.<sup>3</sup> It is therefore entirely possible that a lot of the value of governance work – as a means to wider ends – is lost if the governance achievements in these programmes are not measured. With an increasing orientation towards the results agenda, this could mean that the governance agenda becomes undervalued despite its importance in many developing contexts.

8. **Assess what are the most important indicator links for Value for Money (VFM) assessments.** To this end, more work is needed to unpack those indicators most useful for monitoring and assessing VFM (effectiveness, efficiency and economy). This is outlined in more detail in the complementary VFM Report, and includes:

- **Guidance for theories of change for Governance and Conflict, with further work needed on locating a balanced basket of indicators within the VFM options,** and linking this to programmes' theories of change and to data sources such as logframes, country databases and management information systems. (see VFM Report, recommendation C)
- **Supplementary guidance on baskets of indicators.** There is a probably small piece of work on writing a supplement to the Logframe Guidance on how indicators need to be relevant and robust individually, but hit a set of criteria as a basket. (see VFM Report, recommendation D)
- **Development of comparative measures for economy and efficiency.** It is acknowledged that developing reliable benchmarks (e.g. unit costings) is a prerequisite for conducting VFM assessments as part of the Business Case procedure. In addition, programmes need to demonstrate that the choice of activities and outputs in the programme design has included consideration of the most effective ways to deliver the outcome ('more for the same', or 'more for less'). (see VFM Report, recommendations E and F).

---

<sup>3</sup> This is a finding emerging from a review of the World Bank's portfolio. Source: Presentation by Nick Manning at the OECD GOVNET meeting, 8<sup>th</sup> December 2010, Paris.

## REFERENCES & FURTHER READING

Andrews, M (2008) *The Good Governance Agenda: Beyond Indicators without Theory*, Oxford Development Studies: Vol 36, No 4. Routledge

Arndt, C and Oman, C (2006) *Uses and Abuses of Governance Indicators*, OECD

Arndt, C and Oman, C, (2008) *The Politics of Governance Ratings*  
<http://www.oecd.org/dataoecd/21/17/40037711.pdf>

Centre for Future State (2010) *An Upside-down View of Governance*, IDS

Church, C, & Rogers, M (2006) *Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programmes*. Search For Common Ground, Washington DC. [www.sfcg.com](http://www.sfcg.com)

DFID (2007) *Gender Action Plan*

DFID (2010), *Governance Portfolio Review: A Review of DFID's Governance Portfolio 2004/5-2008/9*, Draft Version

DFID (2009) *How to Note: Guidance on using the revised logical framework* (02/09 version)

DFID (2009) *How to Note: A Practical Guide to Assessing and Monitoring Human Rights in Country Programmes* (09/09 version).

DFID HTN M&E

DFID (2007) *How to note: Country Governance Analysis*

Eyben, R (2010) *Hiding Relations: the irony of 'Effective Aid'*, *European Journal of Development Research*: (1-16).

Leaderach, J-P, Neufeldt, R, Clubersto, H (2007), *Reflective Peacebuilding: a planning, monitoring and learning toolkit*, CRS and Institute for International Peace Studies. [www.crs.org](http://www.crs.org)

Mendizabel, E & Clarke, J (2010) *Monitoring and Evaluation of Policy Influencing Interventions* (08/10 draft version). Rapid for DFID

Moehler, D.C. (2010). *Democracy, Governance, and Randomised Development Assistance*. *The ANNALS of the American Academy of Political and Social Science*, 628 (1):30-46

Natsios, A (2010) *The Clash of Counter-bureaucracy and Development*, Centre for Global Development

Rao, S (2010) *GSDRC Helpdesk Research Report: Evaluation of Governance Programme Indicators*, GSRDRC

Skuse, A (2007) *Monitoring and Evaluating Communication for Development*, DFID, London. <http://www.dfid.gov.pub/files/icd-guidelines.pdf>

Sudders, M & Nahem, J (2004) *'Governance Indicators: A User's Guide*, UNDP Oslo Governance Centre, <http://www.undp.org/oslocentre/docs04/useguide.pdf>

Tripathee, S (2007) *Monitoring a Moving Target: Peace-Building Soap Opera in Nepal*, Search for Common Ground

UNDP (2006) *'Measuring Democratic Governance, A framework for selecting pro-poor and gender sensitive indicators'* UNDP, New York

Williams, A. & Siddique, A (2008), *The use (and abuse) of governance indicators in economics: re review'* *Economics of Governance* 9:131-175

Williams, G (2010), *What makes a good governance indicator?* Policy Practice Brief No 6. Policy Practice.

World Bank (2004), *'Poverty Monitoring Guidance Note 1: Selecting Indicators'*, World Bank, Washington, DC

## Appendix 1: Terms of Reference

### How Do You Measure The Impact & VFM of Governance Programmes? October 2010

#### Introduction

DFID needs to improve the monitoring of both its impact and value for money (VFM) of governance programming. In-house work to date, both on global governance assessments and the 2010 Governance Portfolio Review, has revealed a lack of standardised indicators internationally in any donor's governance activity.

Where governance assessments and/or indicators do exist, they are either

- not very user friendly (for example, USAID's 300-page set of democracy & governance indicators from 1998);

or

- they cannot provide attribution information (for example, the Worldwide Governance Indicators dataset, which monitors country progress over time on various aspects of governance. These are, arguably, the most comprehensive set of indicators of governance performance available, but they do not (seek to) explain the reasons why any country may have improved [or worsened] its governance performance).

As such, a draft list of suggested indicators for use *at the programme level* has been prepared by DFID (see Annex 1), setting out possible indicators of outcome- and output-level activity. This list now needs to be tested and updated / revised accordingly, for use in future DFID programming.

#### Objectives

There are two overall objectives for this piece of work:

- To test the relevance and robustness of the attached draft list of suggested indicators and to assess which are the most suitable for different programming purposes (which indicators best tell us whether we have achieved what we set out to do?). Where indicators are considered inappropriate, alternative suggestions should be provided;
- To set out how value for money can best be measured in governance and conflict programming, and whether the suggested indicators have a role in this or not.

#### Scope

This work applies across the spectrum of governance and conflict programming, including support activities on security & justice; civil service reform; elections; parliamentary strengthening; political party capacity building; the media; empowerment and accountability; anti-corruption; tax / revenue generation; human rights; peacebuilding & peace process support; demobilisation, disarmament & reintegration (DDR); conflict prevention & reconciliation; conflict sensitivity; and addressing the underlying drivers of conflict.

The study will not be looking specifically at indicators on Public Financial Management (PFM) as work on PFM has already been taken forward and agreed internationally. There may be some need to discuss some Public Expenditure and Financial Accountability (PEFA) indicators with DFID's PFM team given the use of some PEFA indicators in other 'themes' of governance, e.g. Civil Service Reform. Any insights on VFM measurement that may apply to PFM reform work should, however, be taken into account to complement ongoing DFID-internal work on this area.

The study will discuss which indicators are most amenable to disaggregation by sex, ethnicity, age and disability.



Next, the work will consider the range of countries in which DFID works, from fragile and conflict-affected states to more stable environments.

Finally, the study will consider whether any sets of indicators across the 'themes' of governance can be aggregated into a 'higher level' measurement of overall governance performance.

It is anticipated that the findings of this work will be used not only by DFID, but also by other donors and counterparts in DFID partner countries. The list of suggested indicators should be finalised, as a guide for any agency undertaking or supporting governance-related reforms in the future.

## Outputs

There are 6 specific outputs from this research:

- A list, no more than 10 pages, of suggested standardised indicators for governance programming - broken down by governance theme, and into outputs and outcomes;
- A more detailed publication of a maximum of 30 pages, setting out the background to the research, methodology undertaken, and providing a narrative explanation for the indicators produced (why chosen; where from); and how they should be used at the country level;
- A presentation at the DFID Global Governance & Conflict Conference in November 2010 for all governance advisers setting out (emerging) findings; this will then be linked to a subsequent presentation of a case study from DFID Nigeria
- An additional day following the Governance & Conflict Conference in November 2010 setting out findings in more detail than possible in the presentation requested above and using 3 existing programme logframes as case studies;
- A lunchtime seminar at DFID for presentation of findings to a wider, non-governance & conflict specialist audience, at a date to be determined;
- Possibly, a presentation at an international donor meeting to set out findings at a date to be determined, likely December 2010.

## Methodology & Budget

There will be two principal stages to this work:

### **Stage 1: Indicator Testing** – desk based exercise

Indicators will be tested for *relevance* through

- **Side-by-side matching** of current output and purpose indicators from selected DFID logframes, with the closest suggested indicators from the draft set
- **Noting:** complete matches, close matches, and non-matches (traffic light system)
- **Tabulation** of the overall level of matching by governance theme
- **Description and analysis of matches and non-matches**
- **Reviewing the quality of the indicators (SMART and SPICED)**
- **Reviewing the project logic** to test whether the indicators are at the right level of the hierarchy
- **Suggestions for revision**

Indicators will then be tested for *robustness* through

- **Elaboration of data sources** in the draft set, with a particular focus on those seen as 'most relevant'
- **Review of sources** through an adapted version of the IMF Data Quality Assessment Framework (DQAF)
- **Tabulation of levels of robustness**

- **Description and analysis** of highest and lowest levels of robustness

### **Stage 2: Measuring Value for Money**

UK National Audit Office (NAO) '3E' framework (Economy, Efficiency, Effectiveness) will be used to assess selected logframes. Work will be combined with other approaches (for example ongoing work with Foreign & Commonwealth Office; social return on investment work within UK NGOs). Focus would start at output and outcome level, and measure both qualitative and quantitative VFM. Draft indicators will then be tested for coherence with proposed approach.

### **Timeframe & Reporting**

The consultants will report direct to Claire Vallings in the Politics & the State Team (PST), Policy Division. An internal DFID reference group will advise on outputs and findings.

This work will be undertaken between mid-October 2010 and end January 2011. A draft report should be submitted to DFID by mid-November 2010, as background reading for the Governance & Conflict Conference.

## **Annex 1 to Terms of Reference:**

### **Suggested *Programme-Level* Indicators for Governance & Conflict Programming**

#### **January 2010**

#### **Background & Introduction**

1. DFID's Results Action Plan sets out 2 principal courses of action that relate to the better monitoring of DFID programme performance: the production of *standard* and then *suggested* programme-level indicators.
2. Standard indicators for use in communications with the UK public were developed during 2009 and must now be *compulsorily* used in all relevant programme logframes. Guidance for their use was published in December 2009.

There are 20 standard DFID indicators, 14 of which monitor progress against activities directly attributable to DFID (output level), with the remaining 6 monitoring purpose or outcome level progress that is not attributable to DFID.

3. None of the standard indicators relates to governance or conflict programme activity. However, the second area of follow up to the Results Action Plan was the production of *suggested* programme-level indicators for all areas of DFID activity. The list below is the first phase of doing so for all governance and conflict activity.

#### **Approach**

4. Three methods were used to draw up this list:
  - Consultation with DFID governance and conflict advisers (GAs / CAs), both from central policy teams and country offices in all 3 of DFID's programmatic regions (Africa; Asia; and the former MECAB);
  - Drawing from existing indicators – such as from the UN Convention Against Corruption (UNCAC) in the case of anti-corruption programmatic work; or Human Rights indicators from the UN Office of the High Commissioner for Human Rights (OHCHR);

- Verification of indicators already being used in ongoing DFID programmes by examining existing logframes.
5. Four 'rules of thumb' were also used whilst drawing up these suggested indicators:
    - Suggested indicators should be comprehensive, to cover the full range of activities in which GAs and CAs may find themselves involved, from Public Financial Management (PFM); to civil society support; to security and justice sector activity;
    - Suggested indicators should be provided to cover the variety of contexts in which DFID works (stable; fragile; conflict affected; as well as regional / geographic considerations);
    - Suggested indicators should be gender- and conflict-sensitive as far as possible;
    - The list should be concise and as 'user-friendly' as possible.
  6. These 'rules of thumb' are inherently contradictory: comprehensiveness does not usually lead to succinct and concise lists. And by attempting to suggest indicators across the variety of contexts in which we work, the number of indicators provided relevant to specific contexts is reduced.
  7. However, experience from USAID – the only other bilateral to have conducted a similar exercise in the past – as well as from the vast industry of governance assessment activity that exists, shows that there is a real danger when embarking on this type of exercise that the result will simply be a huge, long list of indicators that are not subsequently either used in country office programming, or which are not monitored so cease to be of practical use in the longer-term.

#### **Flaws in this (type of) list**

8. It should be recognised from the start that this list is by no means exhaustive, and nor is it supposed to be. The context of each country where DFID advisers are based will vary too much for attempting to provide indicators for every possible intervention in any meaningful or useful way. This is a list of suggestions, and should be treated as such: there is no compulsion to use these indicators, and it is certainly acceptable to alter them to make them more relevant to specific programme design.
9. It should also be recognised that by attempting to improve DFID's measurement of governance and conflict activity and impact, indicators will automatically be largely quantitative. This is not to suggest that only quantitative indicators are a good measurement of governance and conflict programming, or that only quantitative indicators should be used in existing or future logframes. Rather, it is anticipated that some of the suggestions below – or variations of them - will be used, but that GAs and CAs will also (want to) include qualitative (perhaps much more context-specific) indicators to monitor the progress of their programmes.
10. The DFID Results Action Plan sets out other flaws inherent in all attempts at monitoring development activity (i.e. not just on governance & conflict):
  - monitoring outcomes is harder than monitoring either outputs or, in particular, inputs;
  - data quality is often poor;
  - attribution of impact is therefore difficult.

Nonetheless, DFID still needs to improve its impact monitoring, collecting data more rigorously than previously, and disaggregating it appropriately. In all of this, good quality statistics are crucial.

11. So this list should be recognised as 'living'. It should be expected that it will change over time, as we monitor uptake and usage of individual or 'clusters' of indicators. This list will remain in draft for the foreseeable future, whilst further research is undertaken.

### **Why now?**

12. There are two reasons why there is growing demand to better monitor our work. Firstly, DFID recognises that it has not, to date, monitored value for money (VFM).<sup>4</sup> Instead, we score programmes against achievements, a practice that is practical, pragmatic and has so far been acceptable to Her Majesty's Treasury (HMT). However, this method also has the disadvantage of not being based on practical, hard, evidence: by only looking at 'goal' and 'outcome' level achievements of our programmes, the impact of our, DFID-specific, investment cannot readily be gauged. Much of DFID's programme monitoring, therefore, remains subjective, and the National Audit Office (NAO) has asked for improvement. It should be highlighted that this is true of all DFID's programme activity, not just governance and conflict programming.
13. Secondly, in an era of global economic downturn, political demands from Parliament to monitor VFM have increased. It is anticipated that this demand will continue for some time.
14. Finally, there is an element to this work around professional integrity: whilst it is well and readily acknowledged that governance and conflict activity can be difficult to monitor, and certainly that impact attribution is difficult to gauge, as a government department DFID has a responsibility to (be able to) account for its work. The improved use of programme level indicators to our programmes will mean we are better able to evaluate our programmes in the future, and thereby better perform our role as civil servants. Better programme monitoring will provide increased insight into good practice, and optimise future programming.

### **How should these indicators be used?**

15. These indicators have been drawn primarily for use by DFID country office GAs and CAs, to use in programme design and monitoring, and particularly in logframes. The following should be borne in mind when they are being used:
  - These indicators should be used to measure DFID's activity and achievements. In the event of multi-donor funding to a programme, DFID's 'share' of attribution can be derived from its proportionate contribution to the overall programme budget;
  - There may prove to be some overlap between outcome and output indicators, depending on (programme) context. Staff should not feel constrained by where individual indicators are set out below; if an indicator listed as an output below would be more valid in a particular programme as an outcome indicator, it should be used as such (and vice versa): an element of staff discretion should be used. However, caution should be taken not to use the same indicator at both output and outcome level;
  - Some suggested indicators can only be proxy measures of activity. Research by DFID Policy Division is ongoing to test programme theories of change, and therefore to examine which indicators may best demonstrate programmatic impact;
  - Using a mix of indicators will strengthen impact measurement; the list below is not suggesting that any of the indicators can be used in isolation from the others;
  - To make the best use of these indicators, results should be disaggregated as far as possible. Data can be disaggregated by sex; age; ethnicity; level of education;

---

<sup>4</sup> Dec 2009 Investment Committee papers on Value for Money

geography; and/or disability. Subject to the type of programmes, or anticipated objective, it is up to advisers to decide how their data should be disaggregated. Those indicators below that can be disaggregated are flagged with “(disagg)”

- These indicators are presented to spark ideas; they are not intended to replace any indicators or impact measurement systems commonly used already by partner countries. Wherever established mechanisms for tracking reform progress exist, these should (continue to) be used;
- If not using any of the indicators suggested below, your own indicator should simply make it clear what you are measuring. For further support, you will want either to speak to your team Statistical Adviser (or contact Claire Vallings in Policy Division’s Politics & the State Team – [c-vallings@dfid.gov.uk](mailto:c-vallings@dfid.gov.uk); +44 20 7023 0366).

### **Implications and Risks of these indicators**

16. Much governance work (and progress against its support) is difficult to measure. The principal implication of this is that, in those areas where DFID is targeting its efforts, perception of client satisfaction may have to be measured as a proxy for progress made. This means that (time and costs of) feedback surveys - an activity DFID has not undertaken on a large scale to date - will need to be factored into programme design.
17. A further implication is that more and better research on both the appropriateness of this type of proxy measure, and which indicators are the best ‘type’ to measure impact, is needed. This work will be undertaken centrally in DFID, and feed into broader discussion on how best DFID can report its impact.
18. It should be noted that DFID senior management are aware of the potential risks of quantifying our work, particularly on governance and conflict, and thereby ensuring a corporate focus only on what is being measured quantitatively (rather than what is important). There is a commitment to ensure against this by reinforcing the importance of work that is less easily quantified.<sup>5</sup>

### **Future work**

19. There are several streams of work to be undertaken:
  - The verification of which indicators are most being used in governance and conflict programming, to improve on current practice and better evaluate programmes;
  - Further research is needed to test programme theories of change and examine which indicators may best demonstrate impact of these theories;
  - Discussion with external (donor) counterparts for their views, both on this list and on impact monitoring of governance and conflict programming more generally. The appetite for potentially establishing some kind of international ‘Governance Results network’ should also be explored.

---

<sup>5</sup> DFID How To Note ‘Standard Indicators’ (October 2009)

## Appendix 2: Original List of Governance and Conflict Indicators

Updated October 2010

<p><b>Security &amp; Justice Sector</b></p> <ul style="list-style-type: none"> <li>○ Include some CSR indicators (recruitment, inspections, promotion, training [how &amp; what], codes of conduct)</li> <li>○ Include some E&amp;A indicators</li> <li>○ Note these do not cover work with coastguards; customs systems; border / immigration</li> </ul>	Purpose / Outcome	Effectiveness of command & control systems - % targeted security personnel stating systems are effective
		Level political interference - % targeted citizens / CSOs stating no political interference in (civilian oversight of) S&J sector
		Quality of prison conditions (against international standards) – average nos. in cells; prevalence of infectious disease / malnutrition; nos. rapes per capita prison population)
		% citizens satisfied with (formal / informal) S&J performance (disagg)
		Confidence in formal systems - % complaints on ‘less serious’ issues
		% citizens who say they feel safe going out in their neighbourhood at night (disagg)
	Output	Command & control systems in place
		Civilian oversight in legislation
		Ratio military personnel: population size
		Per capita military expenditure (or % GDP)
		% citizens who say they have access to (formal / informal) court systems to resolve disputes (disagg)
		# cases resolved using alternative dispute resolution mechanisms
		% citizens who pay a bribe to access police / judicial services (disagg)
		% police / citizen priority similarities
		Length of pre-trial detention - % inmates awaiting trial
		Length of judgement / appeals processes
		# CSOs consulted on policy development

<b>Civil Service Reform (CSR)</b> <ul style="list-style-type: none"> <li>○ (link to sectors here: S&amp;J; Education; Watsan; Health; Infrastructure)</li> <li>○ Cross check with E&amp;A indicators</li> </ul>	Purpose / Outcome	Transparency of inter-governmental fiscal relations – PEFA Indicator 8 (Central or Sub-National)
		Ratio ghost workers: total staff
		Effectiveness of recruitment / promotion systems - % targeted staff who consider promotion systems fair
		Effectiveness of recruitment / promotion systems – vacancy rate
		Improved govt. statistical capacity - % data on X collected by trained government staff
		Ratio sector budget allocation: expenditure - PEFA Indicator 2
	Output	Civil Service Code in place – # targeted staff aware of civil service code
		Effectiveness of payroll controls – PEFA Indicator 18
		% staff with job description
		% staff trained to do their jobs (admin mgmt; policy dev; financial mgmt)
		% staff who understand how their role leads to frontline service delivery
		% senior staff strategic planning training
		Ratio male: female staff in targeted ministries
		% staff minority representation against national stats
% deadlines met (e.g. salaries / budget)		
% complaints received acted on by Ombudsman		
# disciplinary cases for violations of codes of conduct		

<b>Elections</b> <ul style="list-style-type: none"> <li>○ Cross Check with CSR indicators</li> </ul>	Purpose / Outcome	Free & Fair elections – % external observation reports stating elections have been conducted ‘freely and fairly’
		Participation equal (sex; ethnicity; marginalised)? - % voting age

		registered to vote (disagg)
		% registered voters who vote
		% satisfaction with conduct of the election (disagg)
		Awareness & engagement - % citizens received voter education (on process, rights & responsibilities) (disagg)
		% citizens who feel able to cast their vote without pressure
		% citizens expressing confidence in capacity of police to prevent electoral violence (disagg)
		Dispute Resolution capacity – No. judges trained in electoral dispute resolution
	Output	Legal framework in place (up to international standards)
		Independent supervisory body in place - % electoral body staff trained in their specific role / with job descriptions (disagg)
		Inclusion - % women / minority members represented in parliament
		% women / minority candidates
		% population aware of election (disagg)
		% population aware how to vote (disagg)
		Process - % observers trained (disagg)
		% elections preparations completed on schedule
		# incidents of political violence reported in national media (pre-election period; election day; post election period)
		% electoral appeals concluded
		Ratio national: international observers
		Ratio observers: population size



<b>Parliament Support</b>	Purpose / Outcome	Access to Information - % lobby groups / CSOs stating they <i>do</i> access voting records / debate information / parliamentary information, including assets
		Parliamentary effectiveness - % citizens satisfied with parliamentary performance (disagg)
		Length of passage of legislative reform – # days from legislative submission to ratification
		Representation - % seats in Parliament / local government occupied by women / targeted groups
		Scrutiny quality - PEFA Indicator 27 (on annual budget law)
		Scrutiny quality - PEFA Indicator 28 (on external audit reports)
	Output	% parliamentarians trained in what their role is and how to be effective (disagg)
		% parliamentarians trained in budget procedures, including scrutinising & monitoring
		% parliamentary cttes technically trained (in scrutiny of budget, public funds, service delivery – including S&J sector)
		% parliamentarians complying with asset declaration mechanisms
		% lobby groups / CSOs stating they <i>can</i> access voting records / debate information / parliamentary information, incl. assets

<b>Political Party Support</b>	Purpose / Outcome	% political parties with issue-based manifesto / codes of conduct / audited accounts
		No. policies communicated to general public
		% citizens able to identify policy differences among parties
		Political party financing law in place - % political party accounts publicly available
		% registered political parties with regulations on internal governance (that are observed)

	Output	% women / minority members of executive committees of political parties
		% political parties producing annual plans and budgets
		% political party membership given votes in internal decision-making
		% political party accounts and committee level meeting minutes published (within party / externally)
		% women / targeted groups included in membership of national political parties (against national statistics)
		% citizens aware of right to join political parties

<b>Media</b>	Purpose / Outcome	Quality media law / regulator - % media outlets satisfied with quality of media regulator
		% targeted citizens access to media (disagg)
	Output	Independent media regulator in place
		% journalists understand role / neutrality
		Male: female journalists at national media outlet
		% journalists taking a bribe / paid independently for their work
		% targeted citizens satisfied with media quality (disagg)
		% programming given to minority issues (incl. equality across political parties; gender)
		% media coverage on corruption issues

<b>Tax / Revenue</b>	Purpose / Outcome	% targeted citizens stating tax laws non-discriminatory
		Level of uniformity of tax collection – PEFA Indicator 15
		% citizens aware importance of taxation - PEFA Indicator 13

	Output	% tax authority staff trained (see CSR indicators)
		% citizens stating they have access to tax information
		% business satisfaction with tax system
		% businesses paying bribe to avoid taxation
		Complaints system in place / % citizens stating it is effective

<b>Corruption</b>	Purpose / Outcome	% reported corruption cases investigated by anti corruption body (domestic or international)
		% investigated cases lead to prosecution and/or sanction (domestic or international)
		% respondents saying that recruitment and promotion in government is based on professional criteria
		% respondents (government or members of public) that did not report a corruption case because they (a) did not know where to report it; (b) felt it would not be addressed; and/or (c) did not feel safe reporting it (disagg)
		% targeted citizens satisfied with government anti corruption efforts (disagg)
	Output	Corruption in all its forms as defined by UNCAC is criminalised by law
		Frequency and scale of bribes paid by public (disagg [including by sector / government body])
		% national budget dedicated to anti corruption body
		% senior civil servants that comply with asset declaration mechanism
		% citizens stating they have access to agency info (disagg)
		% increase in assets frozen or returned from corruption cases

<b>Human Rights</b>  ○ Cross check with Corruption; E&A; S&J indicators	Purpose / Outcome	% Universal Periodic Review (UPR) recommendations implemented
		Level of government accountability for service delivery – % objectives delivered by human rights body
		# CSOs (per 100,000 persons) involved in promotion / protection of right to X
	Output	% (targeted) citizens access to (health; education; S&J) services (disagg)
		% citizens aware of right to access to basic services (disagg)
		% citizens believe HR body is improving human rights situation in their country (disagg)
		% national budget targeting health / education / food security etc
		% household budget spent on accessing health / education / Watsan / S&J etc services
		% reported cases of domestic and/or sexual violence responded to (by government)

<b>Empowerment &amp; Accountability</b>	Purpose / Outcome	Legislation in place: % targeted CSOs / community groups / citizens stating (a) awareness of access to info and/or (b) that they have accessed it
		PRSP = Pro Poor? / Participatory budgeting: % relevant CSOs stating they were consulted in PRSP / sector plan design / budget allocation (to an extent to which they are satisfied) (disagg)
	Output	Legislation on statistical publishing exists in place?
		# PPAs used in policy design
		% proposals from consultations used in national strategy documentation (PRSP / budget [national and sub-national] / sector programme)
		% sub-national expenditure covered by PETS
		# CSOs tracking budgets

		% government expenditure online / in public libraries
		% national statistics publicly available (online / on paper outside National Statistical Office)

<b>Peace Process Support</b> <ul style="list-style-type: none"> <li>○ Cross check with Elections, Political Party support</li> </ul>	Purpose / Outcome	% negotiating partners publicly advocating peace
		% targeted citizens content with progress towards peace process
		% of terms of peace agreement successfully implemented
	Output	% targeted citizens aware of peace process
		# ceasefire violations (if ceasefire in place)
		% peace process mediators trained in conflict resolution
		Existence of a well-resourced mediation team
		% stated negotiation issues that align with drivers of conflict/grievances identified by population and warring parties
		% identified resources/capacities for peace agreement implementation in place
		Primary negotiations backed up with viable alternative dialogues and multi-track diplomacy

<b>Demobilisation; Disarmament; Reintegration (DDR)</b>	Purpose / Outcome	% of ex-combatants who feel they are included as members of their communities (disagg)
		% of target population who report positive attitudes to civ-mil relationships, and to reintegrated ex-combatants
		% of ex-combatants who see a viable future for themselves without returning to armed forces
	Output	# ex-combatants disarmed and demobilised (disagg)
		% ex-combatants able to maintain an independent livelihood and support their families (disagg)

		% ex-combatants who report addressing grievances through non-violent channels e.g. government or traditional resolution mechanisms
		# ex-combatants undergoing traditional cleansing or other ceremonies of acceptance into communities
		% of community who do/don't feel threatened by presence of ex-combatants
		Measure of ease of access to weapons / # weapons in community

<b>Conflict prevention and peacebuilding</b>  ○ Cross check with S&J; Elections; Media; E&A	Purpose / Outcome	# conflict deaths
		% / # population displaced (disagg)
		% reported violent incidents where there is an attempt at a non-violent response (e.g. mediation by elders)
	Output	% targeted citizens trained in conflict resolution
		% target populations taking part in reconciliation activities (disagg)
		% civil society activism promoting non-violence and peace
		# months taken to deal with a case through transitional justice system
		% target population expressing satisfaction with transitional justice process
		% citizens stating satisfaction with effectiveness of regional/government/community structures responding to key conflict issues (e.g. land registry, village courts)
		# days taken to mitigate key conflict triggers (e.g. food price hikes, electoral violence)
# stakeholders stating crisis response coordinated effectively (e.g. between diplomatic and development channels, and between different actors)		

<b>Addressing underlying drivers of conflict</b>  <ul style="list-style-type: none"> <li>○ Cross check with: Media; Human Rights; Corruption; E&amp;A</li> </ul>	Purpose / Outcome	Horizontal inequalities identified in analysis addressed, e.g.
		<ul style="list-style-type: none"> <li>○ % target populations (youth, minorities, other vulnerable groups) who feel discriminated against by government and political systems</li> <li>○ % minority access to government employment</li> <li>○ Inclusiveness of governing regime (using Polity IV, or % citizens stating satisfaction)</li> <li>○ % target population with access to services (e.g. improved water sources, hospitals)</li> </ul>
	Output	Grievances and other drivers of conflict identified in analysis addressed, e.g.
		<ul style="list-style-type: none"> <li>○ % unemployment of youth and other risk groups for violent conflict (disagg)</li> <li>○ % citizens stating religious freedom curtailed</li> </ul>
		% target population reporting improved satisfaction with involvement in local political processes (local government response to issues raised, etc.)
	% target population reporting non-discrimination by service providers (disagg)	
	# youth and other risk groups receiving vocational training (disagg)	
	% trained youth/other groups who report obtaining sustained employment/livelihoods (disagg)	

<b>Conflict Sensitivity</b>	Purpose / Outcome	Existence of a shared analysis of conflict amongst stakeholders.
		Development, peacebuilding and conflict prevention programmes reflect conflict analysis.
	Output	% programmes explicitly addressing drivers of conflict identified in conflict analysis
		% programmes regularly carrying out, as a minimum, a Do No Harm-type review and implementing changes to programming as a result
		Mechanisms in place to review and update conflict analysis and response on a regular basis

## Annex 1: PEFA Indicators

<b>A. PFM OUT-TURNS: Credibility of the budget</b>	
PI-1	Aggregate expenditure out-turn compared to original approved budget
PI-2	Composition of expenditure out-turn compared to original approved budget
PI-3	Aggregate revenue out-turn compared to original approved budget
PI-4	Stock and monitoring of expenditure payment arrears
<b>B. KEY CROSS-CUTTING ISSUES: Comprehensiveness and Transparency</b>	
PI-5	Classification of the budget
PI-6	Comprehensiveness of information included in budget documentation
PI-7	Extent of unreported government operations
PI-8	Transparency of inter-governmental fiscal relations
PI-9	Oversight of aggregate fiscal risk from other public sector entities
PI-10	Public access to key fiscal information
<b>C. BUDGET CYCLE</b>	
C (i) Policy-Based Budgeting	
PI-11	Orderliness and participation in the annual budget process
PI-12	Multi-year perspective in fiscal planning, expenditure policy and budgeting
C (ii) Predictability and Control in Budget Execution	
PI-13	Transparency of taxpayer obligations and liabilities
PI-14	Effectiveness of measures for taxpayer registration and tax assessment
PI-15	Effectiveness in collection of tax payments
PI-16	Predictability in the availability of funds for commitment of expenditures
PI-17	Recording and management of cash balances, debt and guarantees
PI-18	Effectiveness of payroll controls
PI-19	Competition, value for money and controls in procurement
PI-20	Effectiveness of internal controls for non-salary expenditure



PI-21	Effectiveness of internal audit
C (iii) Accounting, Recording and Reporting	
PI-22	Timeliness and regularity of accounts reconciliation
PI-23	Availability of information on resources received by service delivery units
PI-24	Quality and timeliness of in-year budget reports
PI-25	Quality and timeliness of annual financial statements
C (iv) External Scrutiny and Audit	
PI-26	Scope, nature and follow-up of external audit
PI-27	Legislative scrutiny of the annual budget law
PI-28	Legislative scrutiny of external audit reports
<b>D. DONOR PRACTICES</b>	
D-1	Predictability of Direct Budget Support
D-2	Financial info provided by donors for budgeting and reporting on programme aid
D-3	Proportion of aid that is managed by use of national procedures

## Appendix 3: REVISED Programme-Level Governance and Conflict Indicators

**Updated January 2011**

This Suggested List of indicators has been primarily drawn up to assist DFID country office Advisors in developing indicators for programme design and monitoring, and particularly in logframes. The list has a number of characteristics:

Firstly, **the list provides a number of suggested indicators, not a standardized set.** There is no compulsion to use these indicators, and the list is by no means exhaustive and nor is it supposed to be. It is certainly acceptable to alter them to make them more relevant to specific programme design, and it should be noted that the thematic sub-sector headings were pre-defined. Some of the sub-sectors (but not all) have had indicators added, but the focus of the revisions has been primarily on adjustments to the existing set based on the review of relevance and robustness (see main report).

Secondly, **this revised list has been quality checked against a number of criteria**, namely: (i) *Clarity*: that indicators are specific and measurable; (ii) *Rule-bound*: that indicators evidence different elements of a result chain, helping practitioners explain a particular theory of change; (iii) *Causally-linked*: that indicators do not exist in isolation but form part of a causal relationship with a subordinate or higher level result; (iv) *Gender and pro poor*: that indicators suites measure changes that might empower women, the poor or other excluded societal groups; (v) *Cross-sectoral*: that indicators suites will contain indicators that resonate across sectors.

And finally **this is a list of indicators, not objectives.** A number of indicators may be used to measure any one objective – and the objectives themselves need to be first defined in relation to the logic and purpose of the intervention (the theory of change). It should also be noted that the CAR framework has deliberately not been made explicit in the list, as this is to avoid overcomplicating the structure – though designers should consider this framework when selecting and developing indicators. There are certainly gaps, particularly in terms of indicators of accountability, and political drivers of change, though there are enough examples that might guide practitioners in different programme contexts.

<b>Security &amp; Justice Sector</b>	Purpose / Outcome	Rating of progress against a joint plan.
		% political appointments made at senior positions in key departments
		Ratio: no. prisoners per no. beds
		% citizens satisfied with police complaints system
		% citizens who say they feel safe going out in their neighbourhood at night (disagg)
		# violent crimes recorded by the police per 100,000 people
	Moved from outputs	# and % cases where S&J legislation is passed without civilian oversight and approval.
		Ratio military personnel to population size
		Military expenditure as % GDP
		% citizens who believe bribes are necessary to access police services
		% inmates awaiting trial
		Length of judgement/appeals process
	Additional indicators	% citizens who are successful in their attempt to secure access to formal/informal government systems (disagg)
		# cases where women's rights are successfully adjudicated
		Progress in implementing a sector wide policy and strategy
		% judicial decisions upheld by higher courts
		% citizens satisfied with cost/quality of legal services provided
		% citizens using primary justice system in last year reporting satisfaction with process

	Output	Rating of joint plan in terms of allocation of responsibility, authority and accountability across sector
		% citizens who say they have access to (formal / informal) court systems to resolve disputes (disagg)
		# CSOs consulted on policy development
		% citizens aware of how to access and use information on justice issues
		# cases where free legal advice has been provided
		% primary justice institutions using systems for recording actions and documenting decisions
		# and % disputes reported to state institutions that are referred to primary justice institutions
		# cases resolved using alternative dispute resolution mechanisms

<b>Civil Service Reform</b>	Purpose / Outcome	Transparency of inter-governmental fiscal relations – PEFA Indicator 8 (Central or Sub-National)  <i>Where no PEFA, then indicators might include:</i>
		<ul style="list-style-type: none"> <li>• % of transfers from central government are determined by a transparent rule-based system (with criteria, formula)</li> <li>• Rating of the timeliness of reliable information on the allocations to be transferred to sub-national government<sup>6</sup></li> <li>• % (by value) of sub-national government expenditure is consistent with central government fiscal reporting (by sector categories)</li> </ul>
		Ratio ghost workers: total staff
		Effectiveness of recruitment / promotion systems:# and % of unfilled posts (Vacancy rate)
		<ul style="list-style-type: none"> <li>• Leavers in the last year as a percentage of the average total staff (Staff turnover)</li> <li>• Percentage of people that are still in post after 12 months</li> </ul>

<sup>6</sup> Such as following the PEFA rating scale: A (“before the start of their detailed budgeting processes”); B (“ahead of completing their budget proposals, so that significant changes to the proposals are still possible”); C (“before the start of the fiscal year, but too late for significant budget changes to be made”); D (“after budgets have been finalized, or earlier issued estimates are not reliable”).

		service
		Ratio sector budget allocation: expenditure - PEFA Indicator 2 <i>Where no PEFA, then indicators might include:</i> <ul style="list-style-type: none"> <li>• % variance between budgeted and actual expenditure (by sector)</li> </ul>
	Output	Effectiveness of payroll controls – PEFA Indicator 18 <i>Where no PEFA, then indicators might include:</i> <ul style="list-style-type: none"> <li>• Rating of degree of integration and reconciliation between personnel records and payroll data<sup>7</sup></li> <li>• Time taken to make required changes to the personnel records and payroll</li> <li>• # of payroll audits undertaken to identify control weaknesses and/or ghost workers in the past 3 years</li> </ul>
		% staff with job description
		% staff trained in the last 12 months to do their job (by job type)
		% complaints received acted on by Ombudsman
		# cases of disciplinary action for violations of codes of conduct per 1,000 staff

<b>Elections</b>	Purpose / Outcome	Free & Fair elections: External observation report conclusions
		Participation equal (sex; ethnicity; marginalised)? - % voting age registered to vote (disagg)
		% voter turnout (disagg)
		% voter satisfaction (with conduct of the election) (disagg)
		% citizens who trust electoral process as means of legitimizing power (disagg)
		# and % political parties satisfied with conduct of election
		# and % political parties who consider electoral commission to be a competent public body
		% elected parliamentarians who consider electoral commission to be a competent public body
		% citizens expressing confidence in capacity of police to prevent and control electoral violence (disagg)
		<b>% citizens who believe courts resolve electoral disputes fairly</b>

<sup>7</sup> Based on PEFA rating categories (using a desk or expert review).

		<b># of internationally standardised electoral legal reforms implemented</b>
		% seats in parliament held by women
		Ethnic or tribal profile of seats in parliament
		Gender profile of parliamentary candidates
		% electoral appeals concluded by courts
		Independent supervisory body in place - % electoral body staff trained to fulfil their specific role and responsibilities (disagg)
		% target citizen awareness of electoral principles and procedures (disagg)
		% population aware how to vote (disagg)
		Process – number and % national election observers trained (disagg)
		Gender profile of polling officials
		National election plan completion rate
		# incidents of political violence reported in national media (pre-election period; election day; post election period)
		Ratio national: international observers

<b>Parliament Support</b>	Purpose / Outcome	Access to Information - % lobby groups / CSOs/ media/ political parties stating they have accessed <del>to</del> access voting records / debate information / parliamentary information, including assets in the past 12 months
		Parliamentary effectiveness - % citizens satisfied with parliamentary performance (disagg)
		Length of passage of legislative reform – # days from legislative submission to ratification # new legislative reform bills ratified in past twelve months
		Representation - % seats in Parliament held by women
		Scrutiny quality - PEFA Indicator 27 (on annual budget law)
		Scrutiny quality - PEFA Indicator 28 (on external audit reports)
		# parliamentarians who declare assets
	Output	# parliamentarians trained (in what their role is and how to be effective) (disagg)

		# parliamentarians trained (in budget procedures, including scrutinising & monitoring)
		# parliamentary cttes technically trained (in scrutiny of budget, public funds, service delivery – including S&J sector)
		# lobby groups / CSOs / media/ political parties who know how to access voting records / debate information / parliamentary information, incl. assets

<b>Political Party Support</b>	Purpose / Outcome	% political parties with issue-based manifesto / codes of conduct / audited accounts
		# and % registered political parties who take actions on legal infringements
		% women / minority members of executive committees of political parties
		% political party membership given votes in internal decision-making
	Output	# and % political parties with budgeted annual plans
		% political parties with accounts available for membership or public scrutiny
		% citizens aware of right to join political parties
		% citizens who are able to mention distinguishing policies of two or more political parties

<b>Media</b>	Purpose / Outcome	Quality media law / regulator - % media owners satisfied with regulatory framework
		% target population who use media as primary source of information (disagg)
		% targeted citizens satisfied with media quality (disagg)
		Gender profile of journalists employed in national media

	Output	Media regulator capability rating
		% journalists who are aware of their responsibilities for neutrality
		Revise and move to outcome
		# of newspapers, television, radio and internet channels, covering issues of inequality and discrimination among state and societal institutions
		# newspapers, television and radio channels, covering corruption, electoral misconduct, political violence stories

Tax / Revenue	Purpose / Outcome	Level of uniformity of tax collection – PEFA Indicator 15 <i>Where no PEFA, then indicators might include:</i>
		<ul style="list-style-type: none"> <li>• % of tax arrears at start of fiscal year which was collected during the fiscal year (average of the last two fiscal years)</li> <li>• Frequency of transfer of the tax revenues to the Treasury (daily, weekly, monthly, or longer)</li> <li>• Frequency of complete reconciliation of tax assessments, collections, arrears and transfers to Treasury</li> </ul>
		% citizens aware importance of taxation - PEFA Indicator 13 <i>Where no PEFA, then indicators might include:</i> <sup>8</sup>
		<ul style="list-style-type: none"> <li>• Rating of clarity and comprehensiveness of tax liabilities (desk/ expert review)</li> <li>• Rating of taxpayer access to information on tax liabilities and administrative procedures (desk/ expert review)</li> <li>• Rating of functional tax appeals mechanism (desk/ expert review)</li> </ul>
		% increase in tax collection, by sector or local government (Tax collection rate)
		Tax in arrears as proportion of tax collected
		% business satisfaction with tax system
% citizens stating complaints system is effective		
	Approval of key legislative changes	

<sup>8</sup> For example, based on the PEFA “A-D” rating scale.



	Output	Complaints system in place
		Development of key legislative changes (Tax legislation)
		Rating based on functional reviews of tax authority and tax offices (Organisational reform & capacity building)
		% increase in number of registered taxpayers (Taxpayer identification and registration)
		Higher proportion of actual collection from detected revenue in tax evasion cases (Tax enforcement)
		Average number of days to complete administrative appeals process (Efficiency measure of appeals system)
		Functional internal audit and inspection system - expert/ desk review (Audit and inspection)
		Rating of transparency of procedures for tax collection (Tax compliance and liabilities)

<b>Corruption</b>	Purpose / Outcome	% reported cases investigated by anti corruption body (domestic or international)
		% investigated cases lead to prosecution (domestic or international)
		% targeted civil service staff that feel safe reporting a corruption case
		% targeted citizens who believe government is committed to tackling corruption in public sector
		% national budget dedicated to anti corruption body
		% senior civil servants/parliamentarians/public office holders that declare assets according to regulations
	Output	# laws stating that corruption is a criminal offence
		% senior civil servants revise and move to outcome
		% citizens who are aware of their right to access to agency info (disagg)
		Appointments to anti corruption body based on competency-based recruitment procedure

<b>Human Rights</b>	Purpose / Outcome	% Universal Periodic Review (UPR) recommendations implemented
		Level of government accountability for service delivery – % targets achieved by human rights body
		Issue profile of CSO human rights advocacy (civil, political, economic, social etc)
		Health, education, security as a % of state budget
		Health, education, security as a % of household budget
	Output	# CSOS (per 100,000 persons) involved in promotion/protection of right to X
		% citizens who are aware government has legal obligation to provide basic education and other essential services for all (disagg)
		% citizens who are aware government has legal obligation to protect them from violence and abuse (disagg)
		# reported cases of domestic and/or sexual violence

<b>Empowerment &amp; Accountability</b>	Purpose / Outcome	PRSP = Pro Poor? / Participatory budgeting: % relevant CSOs stating they were consulted in PRSP / sector plan design / budget allocation (to an extent to which they are satisfied) (disagg) <ul style="list-style-type: none"> <li>• % of targeted CSOs that document an impact on new laws or bills</li> <li>• % of targeted CSOs that document an impact on sector policies</li> <li>• % of targeted CSOs that document an impact on the budget process</li> </ul>
	Output	% of targeted CSOs that confirm they can obtain specified information from key public agencies
		# PPAs used in policy design # documented instances where PPAs are used in policy design
		% proposals from consultations used in national strategy documentation (PRSP / budget [national and sub-national] / sector programme) <ul style="list-style-type: none"> <li>• % of targeted CSOs document adequate consultation in the PRSP process</li> <li>• % of targeted CSOs document adequate consultation in the</li> </ul>

		sector planning process
		<ul style="list-style-type: none"> <li>• % of targeted CSOs document adequate consultation in the budget process</li> </ul>
		% sub-national expenditure covered by PETS
		% of targeted CSOs undertaking budget tracking
		% government expenditure online / in public libraries
		% of all national statistical publications available online

<b>Peace Process Support</b>	Purpose / Outcome	% negotiating partners publicly advocating peace through promotion of next plausible steps in peace process
		% citizens content with progress of peace process
		% terms of peace agreement-implemented successfully according to independent assessment
	Moved from output	# ceasefire violations (if ceasefire in place)
	Output	% targeted citizens aware of peace process
		% peace process mediators trained to international standard in conflict resolution
		% negotiations facilitated by negotiation team
		% drivers of conflict included in negotiations
		Rating of extent to which negotiation plan includes multiple approaches to influencing and a range of workable solutions

<b>Demobilisation; Disarmament; Reintegration (DDR)</b>	Purpose / Outcome	% of ex-combatants who feel they are included as members of their communities (disagg)
		% of target population who report positive attitudes to civ-mil relationships, and to reintegrated ex-combatants
		% of ex-combatants who see a viable future for themselves without returning to armed forces
		% ex-combatants who report addressing grievances through non-violent channels e.g. government or traditional resolution

		mechanisms
		% of community who do/don't feel threatened by presence of ex-combatants
	Output	# ex-combatants disarmed and demobilised in past twelve months (disagg)
		# ex-combatants undergoing traditional cleansing or other ceremonies of acceptance into communities
		Measure of ease of access to weapons / # weapons in community/ # households who own a weapon/unit cost of AK47's/# weapons sellers in local bazaar

<b>Conflict prevention and peacebuilding</b>	Purpose / Outcome	# conflict deaths
		# and duration of ceasefires
		% / # population displaced (disagg)
		% population satisfied that transitional judicial procedure is fair
		% citizens satisfied with community structure (e.g. village courts) response to key local conflict issue (e.g. land registry)
		# stakeholders stating crisis response coordinated effectively (e.g. between diplomatic and development channels, and between different actors)
	Output	% incidents where mediation was successful in avoiding conflict (where risk of escalation to violence had been deemed moderate or high)
		# and % targeted citizens trained in conflict resolution
		% target populations taking part in organised reconciliation events (disagg)

<b>Addressing underlying drivers of conflict</b>	Purpose / Outcome	Horizontal inequalities identified in analysis addressed, e.g. <ul style="list-style-type: none"> <li>o % target populations (youth, minorities, other vulnerable groups) who feel discriminated against by government and political systems</li> </ul>
--------------------------------------------------	-------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

		<ul style="list-style-type: none"> <li>○ % minority access to government employment</li> <li>○ Inclusiveness of governing regime (using Polity IV, or % citizens stating satisfaction)</li> <li>○ % target population with access to services (e.g. improved water sources, hospitals)</li> </ul>
		<p>Grievances and other drivers of conflict identified in analysis addressed, e.g.</p> <ul style="list-style-type: none"> <li>○ % unemployment of youth and other risk groups for violent conflict (disagg)</li> <li>○ % citizens stating religious freedom curtailed</li> </ul>
		% target population reporting no-discrimination by service providers (disagg)
		% target population reporting improved satisfaction with involvement in local political processes (local government response to issues raised, etc.)
	Output	# youth and other risk groups receiving vocational training (disagg)
		# and % trained youth/other groups who report obtaining sustained employment/livelihoods (disagg)

<b>Conflict Sensitivity</b>	Purpose / Outcome	Existence of a shared analysis of conflict amongst stakeholders.
		Development, peacebuilding and conflict prevention programmes reflect broader interpretation of root causes of conflict .
	Output	% programmes explicitly addressing drivers of conflict identified in conflict analysis
		% programmes regularly carrying out, as a minimum, a Do No Harm-type review and implementing changes to programming as a result
		# programme reviews and revisions undertaken in the past twelve months

## Making good use of the indicator list

The experience from USAID – the only other bilateral to have conducted a similar exercise in the past – as well as from the vast industry of governance assessment activity that exists, shows that there is a real danger when embarking on this type of list - and that the result will simply be a huge, long list of indicators that are not subsequently either used in country office programming, or which are not monitored so cease to be of practical use in the longer-term. For this reason, this list is more modest in its length and far from being exhaustive. Below are some practical steps for making the best use of the list:

- First, it is essential that the chosen suite of indicators relates to the particularities of the intervention's objectives – and that **the analysis of theory of change should precede the selection of indicators**. The list does not attempt to incorporate an implicit theory of change for each sub-sector. In developing indicators, staff should refer to the 'How To Note' on the Logical Framework as well as other guidance on identifying causal logics, developing objectives and corresponding assumptions.
- Second, **it is essential that the selected indicators capture the theory of change (what really counts), while also lending itself to data capture (what is really countable)**. Do not be constrained by the list. It is perfectly acceptable to take indicators from several thematic sub-sectors, as the sub-sectors are loosely based on the current DFID governance portfolio – rather than being drawn from a thorough analysis of the portfolio or a particular theoretical framework of governance. Staff should use common sense and the myriad of guidance available when developing indicators. In essence, the indicators are presented here are designed to spark ideas; they are not intended to replace any indicators or measurement systems that are already commonly used by partner countries. Wherever established mechanisms for tracking reform progress exist, these should (continue to) be used.
- Third, **staff should consider the power balances/ imbalances associated with the theory of change and reflect on the indicators that have been selected**. This will help ensure that the chosen indicators enable the gathering of information of relevance to the power balance/ shifts in the proposed theory of change (including gender and pro-poor dynamics).
- Fourth, **it is important that a basket of indicators provides a balanced set**. We recommend moving away from SMART-type criteria that tends to focus on the quality of an *individual indicator*, to an approach that leads to a *multi-dimensional basket* of mixed-typology indicators. It is important that the whole set of indicators adds up to more than the sum of its parts. This can be achieved through balancing for example: quantitative and qualitative approaches; technical and political, including some proxy indicators; triangulating perception-based indicators with more fact-based measures; capturing the views of multiple interest groups, and not defaulting to disaggregation to measure change in gender or pro-poor power relations.
- Finally, **review the final basket of indicators to see whether information gathered against them would tell a story** that would confirm/ correct your theory of change.