Measuring the Impact and Value for Money of Governance & Conflict Programmes

Final Report

December 2010

Submitted by

ITAD

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

Contents

List of Figures

1 Introduction

In October 2010, ITAD was commissioned by the DFID Politics and the State team to conduct research and propose a way forward for Governance programmes in conducting value for money assessments as part of a consultancy on measuring the impact and value for money of DFID Governance programmes. The specific objective stated for our work on value for money (VFM) in the Terms of Reference was:[1]

> *"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".*

This objective was taken to involve three core tasks: first, developing a value for money approach that applies to both the full spectrum of governance programmes, and those programmes undertaken in conflict-affected and failed or failing states; second, that the role of a set of suggested indicators should be explored and examined for their utility in this approach, and, further, that existing value for money frameworks (such as the National Audit Office's use of the 3Es of 'economy, efficiency and effectiveness') should be incorporated, as outlined in the Terms of Reference.

As an initial step, ITAD collated information on previous work in DFID on value for money, and then supplemented this with further research on the application of VFM across Government Departments[2] and in other agencies. Current and emerging work on VFM, such as the work of the Audit Commission and work on Social Return on Investment (SROI) was also referenced.

Building on the general principles derived from this exercise, two outputs were produced. First, a conceptual framework was developed to locate and understand different components of VFM in relation to one another, defining key terms and the relationships between managing and measuring value for money. Second, to support putting VFM into practice, a range of possible options for improving VFM through specific actions, starting with results-based management and moving towards explicitly measuring and comparing costs and benefits, were identified. These options include incorporating indicators into VFM assessments.

This Report presents background on VFM from documentary research (section 2); explains the analytical framework that captures key concepts in VFM, and sets out options for improving VFM (section 3). It outlines one specific option, a "3 Es ratings and weightings approach to VFM" as presented to Governance and Conflict Advisers at a DFID Research Day on 25 November 2010, and includes their response plus some initial reactions from Finance and Corporate Performance Division (FCPD), particularly

---

[1] The second objective of the TOR (i.e. to test the relevance and robustness of the Suggested Indicator list) is covered in another complementary report.

[2] Including lessons learned from recent attempts to devise indicator based approaches to impact assessment and improved resource allocation, led by the Foreign and Commonwealth Office as part of the cross-Whitehall Prevent and CBRN Security Programmes.

with regard to Business Case compatibility (section 4). Finally, the Report proposes ways in which initial findings can be refined and further developed to support Governance programming and build staff competence and confidence in conducting VFM assessments (section 5).

2 Background and findings from research

- **What is value for money, and how is DFID taking VFM forward?**
- **Key findings from review of value for money approaches**
- **Economy, efficiency and effectiveness**
- **The role of indicators in value for money assessments**

What is value for money, and how is DFID taking VFM forward?

'Value for money' is a term generally used to describe an explicit commitment to ensuring the best results possible are obtained from the money spent. In the UK Government, use of this term reflects a concern for more transparency and accountability in spending public funds, and for obtaining the maximum benefit from the resources available. At a time of reduced public expenditure and rationalization of resources, this term has gained wide currency in the formulation of economic policy imperatives. It is part of the Structural Reform agenda for DFID, and forms part of wider organizational changes initiated in DFID following the election of the Coalition Government in May 2010[3].

In parallel with ITAD's work on VFM, the Finance and Corporate Performance Department (FCPD) were tasked with developing a new Operational Planning process and a Business Case procedure to support the changes and to bring DFID programme planning and appraisal in line with Treasury requirements for all Government Departments. On appointment, the new Secretary of State for DFID, Andrew Mitchell, emphasised the importance of this work in his first public statement in which he commented that:

> *"Our bargain with taxpayers is this: in return for contributing your hard-earned money to helping the world's poorest people, it is our duty to spend every penny of aid effectively. My top priority will be to secure maximum value for money in aid through greater transparency, rigorous independent evaluation and an unremitting focus on results."*

This commitment continues to be given the highest priority, both internally and through external scrutiny of DFID, for example through the Parliamentary Public Accounts Committee's reviews of Departmental expenditure and results – and the forthcoming

---

[3] Key changes: Action 2.4.iv: Introduce new project design, appraisal and monitoring templates to ensure greater focus on results, risks and value for money. *Draft Structural Reform Plan*, DFID, July 27 2010; *UK Multilateral Aid Review and Bilateral Aid Reviews* due to report in February 2011; the introduction of the new *Aid Transparency Guarantee*; creation of a new *independent Aid Watchdog*.

Independent Commission for Aid Impact that will report to the International Development Select Committee.

VFM practices and approaches: key findings from research

VFM is a set of assessment practices for appraisal, review or evaluation of systems and functions as well as initiatives, schemes and projects that are time bound. The development of these practices, following economic policy imperatives, has traditionally been led by the auditing and accountancy professions. VFM is now found not only amongst a range of internal and external controls, procedures and assessments for financial and resource management, (for example risk management, fraud and corruption, audits[4]), but also forms part of performance management systems and processes in organizations. Hence there are many and necessary links to results-based management, monitoring activity, impact assessment and evaluations.

VFM has expanded well beyond early work on economy savings to now encompass a much more comprehensive approach including outcomes, e.g. recent work has been done by NAO on developing a more comprehensive analytical framework for auditing VFM[5]. VFM is now used by donors and multilateral agencies as part of their appraisal and evaluation procedures (e.g. USAID, DANIDA, Millennium Challenge Commission ERR analysis), and there is renewed interest in its application to aid and development programmes (e.g. World Bank, OECD).

In summary, there are broadly 2 forms of VFM assessment in use: first **auditing of performance management and measurement systems** (by independent UK audit agencies such as NAO, Audit Commission) and second **economic appraisal** (as outlined in HM Treasury's Green Book and adapted by Government Departments). The former entails conducting regular (usually annual) audits of departmental systems, judging performance in key functions by using set criteria and evidence, and arriving at an overall VFM judgement by comparing actual with planned performance and by using external benchmarks, e.g. against accepted good practice, trend analysis or alternative actions such as those employed by similar agencies.

The latter entails using an appraisal process (usually formalised in a procedural template) to decide whether to invest in a proposed scheme or project. This is commonly deployed as a procedure for allocating resources, and is based on assessing net present value (whether expected discounted additional benefits outweigh costs), and comparing various options, including the 'do nothing' option, to select the one that offers the best return on investment.

One obvious contrast in these approaches has been that auditing focuses more on examining performance management processes to determine whether value for money has been secured (i.e. if the procedures are warranted, then value for money should result), whilst appraisal directly measures and compares costs and benefits to determine

---

[4] Examples include the Audit Commission's Use of Resources Framework, CIPFA

[5] NAO Analytical framework for assessing Value for Money.

value for money. However, the expansion referred to above is leading to increased use of more precise measurement of costs and benefits *of systems*, as exemplified in recent work by the NAO.

A Key Approach: Economy, Efficiency and Effectiveness

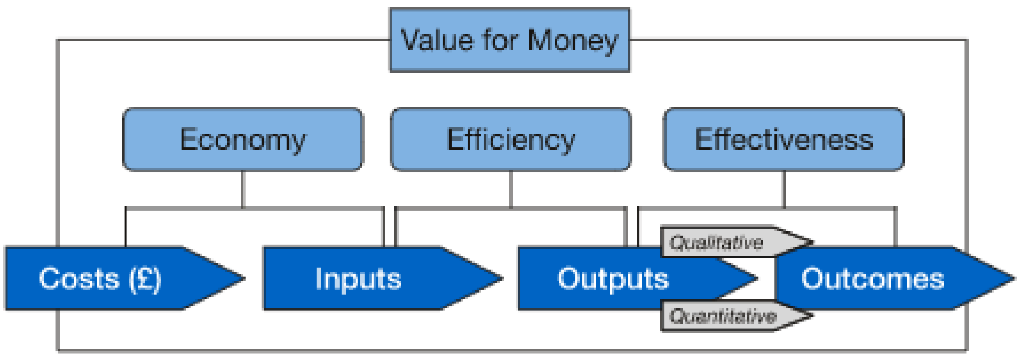
As a starting point, ITAD adopted the NAO '3E' framework of 'Economy, Efficiency and Effectiveness', which follows the same categorisation as the Audit Commission's definition and approach to VFM. Their definition of value for money is:

> *"VFM is about obtaining the maximum benefit over time with the resources available. It is about achieving the right local balance between economy, efficiency and effectiveness, or, spending less, spending well and spending wisely to achieve local priorities...VFM is high when there is an optimum balance between all three elements, when costs are relatively low, productivity is high and successful outcomes have been achieved."*[6]

Further to this definition, the NAO have produced the following definition and analytical framework that states that: "*Good value for money is the optimal use of resources to achieve the intended outcome*".[7]

The assessment of VFM thus involves examining each of the 3 elements of VFM, identifying the links between them and drawing conclusions based on evidence about how well they perform together. Interestingly, the definitions also refer to an optimal balance, as contrasted with a 'maximum' productivity ratio, suggesting that it is not the case that the cheapest option always represents better value for money, and pointing to the conversion of inputs-outputs and outputs-outcomes as the subject of real interest in value for money judgements.

Figure 1: Defining Value for Money



In this model, the following definitions of elements apply:[8]

- **Economy**: 'a measure of what goes into providing a service'. This costs inputs. Unit costs are typically used as an economy measure. 'The whole life costs of

[6] Taken from Audit Commission webpage, accessed 12 October 2010

[7] NAO Analytical Framework, op cit.

[8] Adapted from *Value for money in the use of resources, 3.5*, Audit Commission webpage, accessed 12 October 2010; direct quotes marked.

inputs such as the direct and indirect costs of acquiring, running and disposing of assets or resources should be considered'.

- **Efficiency**: 'a measure of productivity, in other words how much you get out in relation to what is put in'. This examines the relationship between inputs and outputs; for example, planned versus actual delivery of milestones by service providers, or benchmarked comparison among programmes working to same or similar outcomes but using different pathways to achieve intended outcomes.

- **Effectiveness**: Qualitative and quantitative measures of increase or decrease in outcomes that show that a programme 'is effective in delivering its intended objectives'. This examines the relationship between outputs and outcomes.

The role of indicators in value for money assessments

These definitions suggest a significant role for indicators in VFM assessment since they provide 'a measure of productivity' (efficiency) and 'qualitative and quantitative measures of increase or decrease in outcomes' (effectiveness). Clearly, then, the quality of these indicators and the accessibility of data to support measurement of progress against them have important implications for VFM.

3 Value for Money Conceptual Framework and Options

- **Initial Considerations**
- **VFM conceptual framework**
- **Options for VFM improvement and assessment**

Initial Considerations

Introducing value for money parameters to the appraisal and impact evaluation of DFID programmes will have a significant effect on programme design, implementation and evaluation. Various documents produced by DFID testify to this[9]. The development of a Governance programme requires sufficient range and depth in the context analysis (social, economic and political) to formulate relevant responses to very different circumstances, and as a consequence Governance programmes display necessarily diverse and complex pathways to successful outcomes[10]. Building strong business cases for investment which respond constructively and realistically to the focus on managing and measuring for results, and obtaining 'more for the same' or 'more for less' is a particular challenge.

These challenges require a broad framework which could be applied to any Governance programme, in any context, and one which could be adapted to reflect differences in how programme outcomes can be valued.

[9] E.g. Interim Guidance Note on Managing and Measuring for Results in FCAS, 2010; VfM and DFID Nigeria Programmes Strategy, Action Plan and How To, A. Gerry, April 2010, VFM 2009 Spot check, (Mar 2010).

[10] The Draft Governance Portfolio Review, October 2010, presents an outline of the difficulties of measuring VFM across the diversity of Governance programmes, (Executive Summary pp 8-9).

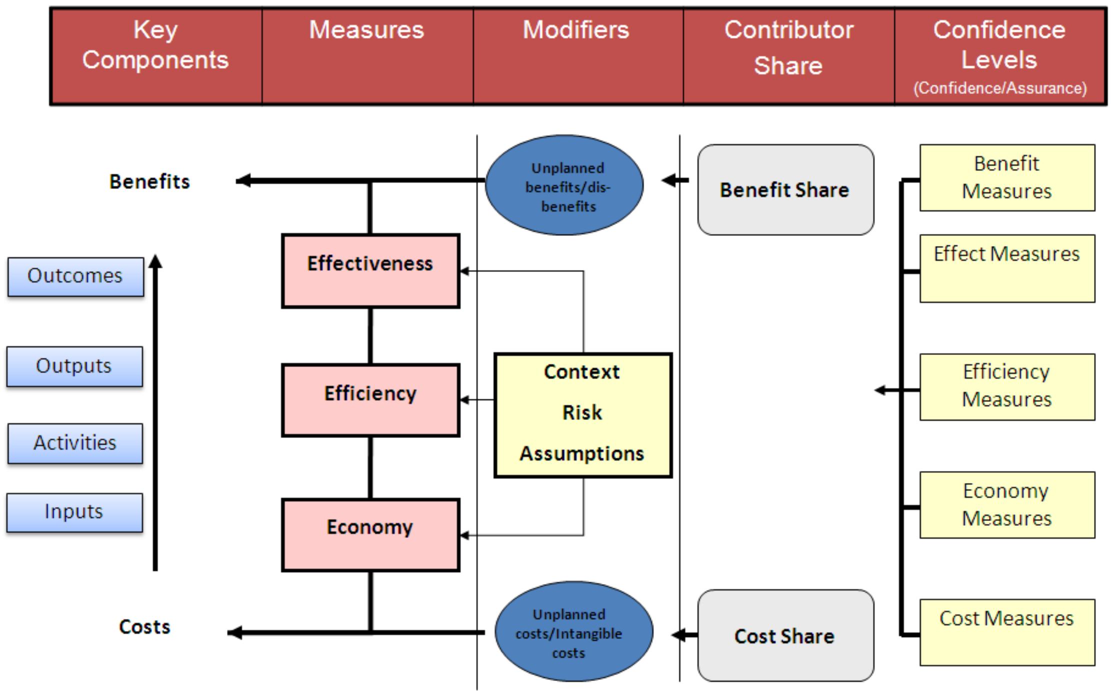
ITAD conducted initial research on the background of previous work in DFID on value for money[11] and supplemented this by further reviewing the application of VFM across Government Departments, (HM Treasury's *Public sector business cases using the Five Case Model: a Toolkit; DfES Evaluation and Appraisal Guidance*), and the work of the Audit Commission (*Use of Resources Framework*, October 2009), as well as the SROI framework[12]. From these general principles, a conceptual framework was developed to locate and understand different elements of VFM in relation to one another, and a range of possible options for improving VFM through specific actions, starting with results-based management and moving towards explicitly measuring VFM, were identified.

These were presented to DFID staff at a meeting on 9th November for discussion and review prior to the Governance and Conflict Advisers' Conference on 23-25th November, and then further developed for a presentation and an introductory exercise to outline the suggested approach to VFM. It was stressed throughout this process, that the primary needs of Advisers were for information, practical assistance and increased confidence in carrying out VFM assessments to support the introduction of the new Business Case and Operational Planning procedures being implemented by DFID from January 2011.

A VFM conceptual framework

To help elaborate and illuminate value for money as a concept and approach, a conceptual framework was developed to bring together and present visually the key elements and considerations (and cautions) of a VFM assessment. The framework is presented and described below:

Figure 2: Value for Money Conceptual Framework

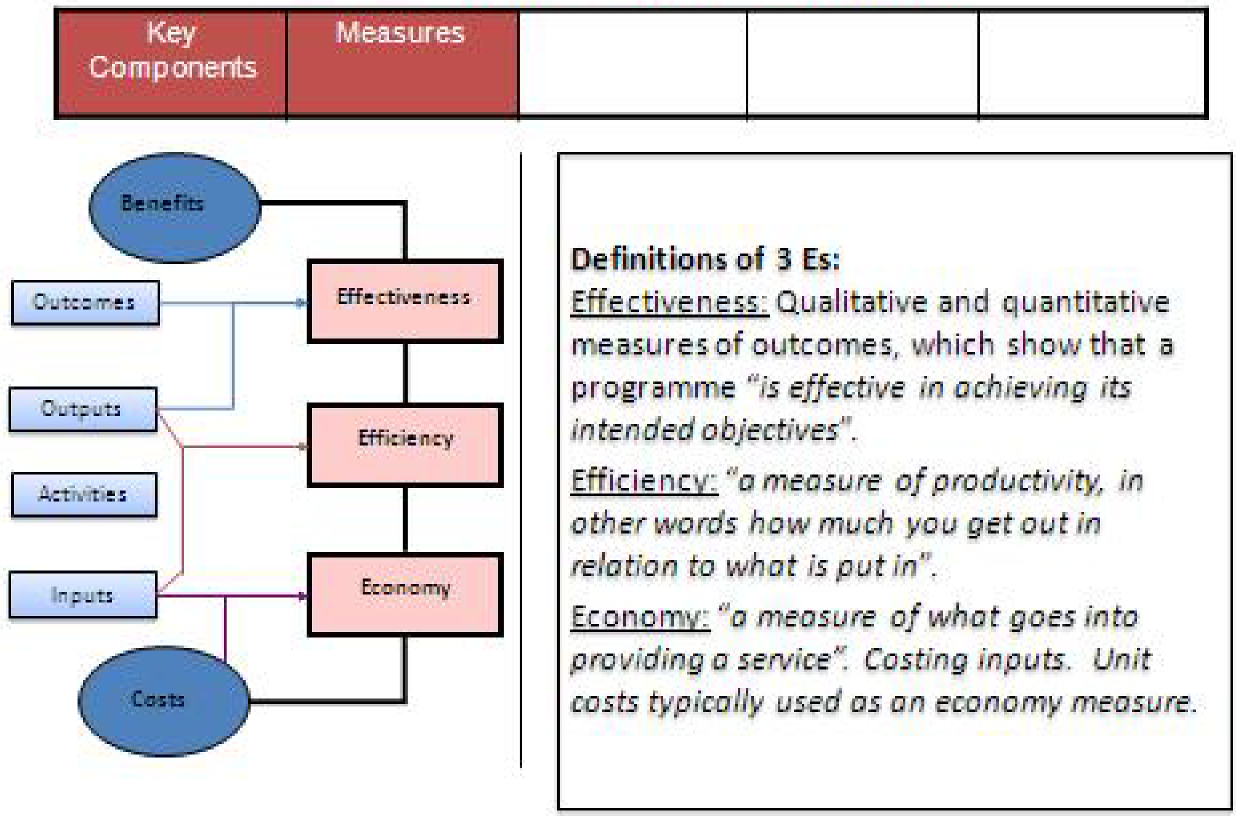


[11] For a list of documents reviewed see Appendix 2: List of reviewed documents and website materials.
[12] *A Guide to SROI*, developed by the Office of the Third Sector, NEF, CES, NCVO, NPC.

1. **Key Components**: Value for money examines the optimal relationship between costs/resources and benefits/outcomes - delivered through processes that transform inputs through activities to outputs which are necessary and sufficient to trigger outcomes (as per the earlier diagram, Figure 1).

2. **Measures**: VFM can be optimized through consideration and strengthening of economy, efficiency and effectiveness processes and measures – to ensure we spend less when we can, spend well and spend wisely (
Figure **2**). It is the balance between these three measures, and not the absolute level of each of them, that represents the optimal route to good VFM. This is a significant point since it implies that in difficult contexts, a reasonable effect may only be realizable at a relatively high input cost (economy). Note below how these 3Es measure the relationship between key levels in a hierarchy of programme objectives:

Figure 3: Key Components and Measures of VFM



3**. Modifiers**: The optimal balance of the 3Es requires the factoring in of context, risk and assumptions which set limits on effectiveness, efficiency and economy. In terms of modifiers, there may also be intangible costs and benefits to factor in which influence judgements on VFM. In addition, at the end of any programme, VFM judgements made at the outset would need to consider not only performance against plan but unplanned costs and benefits.

4. **Contributor share**: There are significant challenges to meet in determining how to attribute costs and benefits when making value for money judgements. This component of the VFM framework is a reminder that the attribution/contribution questions need to be answered in any VFM judgement and that assuming a pro-rata claim on outcomes based on inputs may be too crude an approach.

5. **Confidence levels**: Data quality plays a significant part in any VFM judgement – this is in part linked to how explicit we are about the assumptions we make about reliability, relevance and robustness of the data sets we used (this relates strongly to indicators) and more explicitly is linked to how sensitive our VFM findings would be to changes in any assumptions made. This latter point is of specific relevance to the options for measuring VFM in a governance context as explained in the options analysis below.

Bearing these considerations in mind, a number of options were considered with a view to judging them in terms of:

- How well they strengthen the 'key components' in the framework
- To what extent they capture and enhance measures of economy, efficiency and effectiveness
- How well they tolerate contextual risk and assumptions
- Whether they deal with attribution questions robustly
- How susceptible they are to poor data quality or associated data assumptions

Box 1: Putting the VFM framework into practice

An example of how this works in practice is shown below. This is only illustrative and for the purpose of showing how the framework raises issues about the different aspects of VFM. For a programme that aims to improve access to justice through the training of judges, then:

- **Economy** covers aspects such as the procurement of trainers, the cost of hiring room, and the cost of hiring trainers – including the possibility of unit costs compared to benchmarks for similar training available in the country.

- **Efficiency:** How many judges are trained? How many trainers did we need? Can a measure of productivity be devised?

- **Effectiveness:** Whether judges trained led to improved access to justice? Are there measurable changes at the outcome level?

- **Modifiers:** Contextual matters, such as it being expensive to run training courses in Afghanistan; Considerations of risk, such as perhaps we may be wrongly assuming that the problem is solved through training.

- **Contributor share:** Even if perception surveys show that people feel that the judicial system is fairer, how do we know that that was down to the training? How do we consider issues of contribution and attribution?

- **Confidence:** How representative was the sample survey? How reliable is the data?

Options for VFM Improvement and Assessment

Based on the VFM principles drawn from background reading and building on the analytical framework above, a model presenting four options for improving and assessing value for money was developed:

1. **Option A: Improved RBM:** in which improvements in results based management and other project management processes are used to demonstrate improved VFM
2. **Option B: 3Es Rating and Weightings Approach:** – in which key processes associated with - and measures of - Economy, Efficiency and Effectiveness are identified and used to rate programmes; and where these processes and measures can be weighted to reflect their relative importance.
3. **Option C: Trends Analysis**: in which a comparison of progress against indicators is used to measure effectiveness and in which baselines, targets and underlying trends are all taken into consideration.
4. **Option D: Cost Benefit Analysis (or SROI):** in which outcomes are monetized/financial proxies are introduced in order to make a direct comparison between costs and the financial value of benefits.

Note that these options are cumulative, building upon each other and progressing from an ability to demonstrate that we **manage** what we do with VFM in mind to an ability to **measure** (and compare) the value we generate with the money we spend.

Figure 4: Options for VFM: from management to measurement

Management — Measurement

Option A: RBM Strengthening	Option B: RBM Rating and Weighting	Option C: Trends Analysis	Option D: Potential SROI (Cost-benefit analysis)
The Management Option: Improving management processes and procedures e.g. planning, procurement, implementation (contract management), monitoring and evaluation).	**Balance of weighted, rated effects to rated and weighted costs.** Develop outcome criteria Develop costing criteria Develop ratings scale Develop weightings against outcome and cost criteria Decide between scoring systems and narrative analysis for VFM judgements.	**Adding trend information** to determine difference made by programmes compared with 'do nothing' scenarios Develop outcome statements based on indicators Develop rating scale against outcome statements Include trends criteria within benefits table System requires baseline info against indicators	**SROI** develops qualitative further into quantitative measures, outcomes monetized and compared with costs Processes involving others – stakeholders and wider Governance community Development of financial proxies Information on underlying trends as well as baseline info Sensitivity to assumptions a major issue

The four approaches are described below in terms of methodology, strengths and limitations, using the framework as a reference point for this analysis. Options A and B were the subject of an exercise undertaken with Governance and Conflict Advisers, the results of which are set out in section 4 of the report. Recommendations for developing and testing these methodologies are summarized in the concluding section of the report.

Option A: Improved RBM

Method: Based on audit practice, this option assumes that good practice leads to good results. Very similar to quality programming or 'effective programming' approaches.

Strengths: Very 'do-able': supports current focus on results and encourages strengthening of existing approaches; strong evidence that improved RBM does lead to improved results.

Limitations: The focus is on the *process* rather than the *product*. Unlikely to be sufficient to satisfy current pressures to attain VFM within DFID.

Option B: 3Es Rating and Weightings Approach

Method: This enhances Option A by selecting a number of management criteria and judging performance against a standard so that programmes can be rated on this basis. Since criteria would vary in importance, they might also need to be weighted. Processes that support management as well as measurements of the 3Es would be used. The use of weightings can help overcome commonly-held assumptions that the VFM is solely about cost-cutting, with a rebalance towards key aspects such as effectiveness.

Strengths: Utilises an auditing framework; introduces possibility of benchmarking standards.

Limitations: Benchmarks not yet developed; contextual differences might make comparison difficult.

Option C: Trends Analysis

Method: This approach explores the scope for using indicators to establish a scale against which progress against outcomes can be measured and compared (at least against the 'do-nothing' option). Requires relevant and robust indicators, baseline information, monitoring against indicators, and an ability to determine contribution.

Strengths: Trends analysis can be used to show a positive contribution of an intervention to a generally downward trend, i.e. that the downward trend would have been greater without the programme. This can be important in many development contexts, particularly perhaps in fragile states.

Limitations: Caution is needed before making cross-country comparisons; requires audit scrutiny to ensure measurement procedures are adhered to.

Option D: Cost-Benefit Analysis

Method: Fully monetized indicators of benefit are compared to costs; involves the development of financial proxies for all programme benefits.

Strengths: This approach provides a simple, comparable read out on the return on investment.

Limitations: The approach is extremely sensitive to assumptions and there is a risk that these are manipulated to generate desirable results.

4 Developing and using criteria to rate and weight programmes

- **Developing criteria for rating programmes (how they were developed)**
- **Rating scale, weighting and scoring programmes against criteria**
- **Table of criteria**
- **Testing a prototype: key points from feedback at VFM Research Day**

Developing criteria for rating programmes

A table of criteria for assessing programmes was developed to support the first 2 options (improving results-based management, and rating and weightings). These were identified as the most useful starting point by DFID staff at the meeting on 9 November – and could be adapted to use *ex ante* or *ex post*. The table follows the 3Es in setting out 3 criteria for effectiveness, 2 for efficiency and 2 for economy. Both aspects of VFM were covered across the 3Es: i.e. supporting better results management through assuring the quality of management *processes*, and supporting more direct *measurement* of the conversion of inputs-outputs and outputs-outcomes.

The table represents initial work and is not comprehensive, and it is not intended to suggest that it meets a 'necessary and sufficient' rule for such criteria. Rather, it presents an example of how programme assessments can begin moving towards more explicit measurements, which support VFM judgements through a combination of better results-based management and the measurement of the relationships between inputs-outputs and outputs-outcomes.

The table integrates qualitative and quantitative assessment and represents an important development towards measuring outputs-outcomes, and inputs-outputs, which can be developed further towards VFM appraisal.

The challenge throughout is to find a way to combine the qualitative assessments (which perhaps have more applicability for Governance programmes) with quantitative measures in a way that is logical and coherent, and will be of practical assistance to those preparing a VFM case.

Effectiveness criteria

The key questions that are asked about programme outputs and outcomes tend to be expressed in simple summative forms, such as 'how many'? 'how much'? and 'how long'? One criterion was developed for 'how many/how much' covering **leverage and replication**. These additional benefits relate specifically to Governance programmes because an expected outcome in capacity, accountability and /or responsiveness identified within a programme can often produce spin-offs, multiplier or replication effects, and leverage further benefits through advocacy and the success of a good example. For example, the outcome of a pilot programme can leverage further funding, or extension of benefits to a wider group, or transfer of learning to other programmes and initiatives. Leverage is relevant to both programmes that are about improving governance directly and programmes that include governance as a key element, such as health or education programming. These additional benefits should be factored into valuing outcomes on the basis that the strength of the evidence provided indicates that it is reasonable to conclude that they will occur.

Two further criteria address the **theory of change** used in a programme and the strength of the **indicators** identified. Both of these criteria strengthen results-based management, providing a test of the assumptions made, how far outputs can be expected to deliver the purpose, and the relevance and robustness of the indicators. The latter incorporates work done on testing the Suggested Indicator List, using the checklist developed.

Duration of benefit(s) – i.e. the length of time a benefit is expected to last - was not developed. This was due to time, but it remains an important criterion.

Efficiency criteria

Two criteria were developed on **productivity** and **risk analysis and mitigation**. Productivity (the ratio between inputs and outputs) is a key VFM measure, and is particularly hard to estimate accurately for Governance programmes. This criterion incorporates activity or output costs relative to similar programmes using benchmarking, asks how the value of outputs is being optimized (e.g. by increasing the proportion of outputs or decreasing the proportion of inputs) and whether the scheduling of activities indicates timely delivery of outputs and allows a more formal measure of productivity through comparing actual with planned milestones.

Risk analysis and mitigation addresses the strength of the conflict and risk analyses conducted for the programme, and the steps taken to mitigate and manage risk. To avoid creating risk averse incentives, it focuses on balancing inherent risk with returns expected, and includes responsiveness to changes in significantly difficult circumstances. This is important because DFID has committed to working in fragile and conflict-affected environments, and VFM does not necessarily imply working in a risk-adverse manner (i.e. a high risk, high impact intervention may also represent sound VFM).

Options criteria

Due to time constraints, this was not developed. It is however a useful criteria that considers whether an options analysis has been undertaken to compare different ways of delivering the same outcomes – and which option provides the best value for money (say, compared to the counterfactual).

Economy criteria

These were developed for **procurement** and for **unit costs**. For procurement, the management of costs and achieving significant reductions were identified alongside the ongoing monitoring of costs, and the effects of economies on outputs and outcomes. Unit costs address 'costs relative to benchmarked unit costs', and any additional benefits identified to achieve a good return on investment. In many of the contexts where Governance programmes operate it may not be appropriate to compare unit costs with internationally agreed standards or between different country contexts. Instead a more reasonable benchmark might be against the unit costs of other donors working in the same sector/ country.

Rating scale

For each criterion, a rating scale of "1" (low) to "5" (high) was chosen to provide an ordinal scale of measurement against which programmes can be scored. Descriptors are particularly important and have been developed to:

- Set standards of performance as a basis for making qualitative assessments of programmes
- Promote consistent and transparent rating by assessors
- Enable continual adjustment of standards over time

Within each point on the scale, the descriptors set out specific characteristics of the criterion that indicate the level of performance, graduating from not meeting the standard to reaching a standard of excellence. For each criterion, a summary statement is provided which summarises the level of performance of the programme.

In addition, some aspects of assessment are deemed to be crosscutting. Risk, for example, runs through 2 criteria as well as being a criterion in its own right, and the strength or otherwise of supporting evidence is used throughout.

Weighting and scoring programmes

For the purposes of introducing this table and testing its applicability with Governance Advisers, a weighting system was developed which allowed scores to be allocated up to a maximum of 100 points, distributing the weighting as evenly as possible between different criteria. This would not be recommended in practice. Further work is needed to both develop the criteria and to provide a system of weighting different criteria to reflect the value placed on different aspects of programme design.

Figure 5: VFM Scoring Sheet

VFM Dimension	VFM Criteria	Score Descriptors 1	2	3	4	5	Score
Effectiveness	**Leverage/ Replication** (Weighting X 3)	• No leverage or wider effects identified • No or very low potential for additional benefits (e.g. scale-up, multiplier or replication) identified	• Some leverage of other activities /investment and wider effects identified • Limited potential for additional benefits (e.g. scale-up, multiplier or replication) identified	• Leverage of other activities / investments described and supported by some evidence • Some potential for additional benefits (e.g. scale-up, multiplier or replication) identified	• Leverage of other activities / investments and wider effects described, and supported by strong evidence • Considerable potential for additional benefits (e.g. scale-up, multiplier or replication) identified	• Leverage of other activities /investments and wider effects described with evidence that shows significant potential for expansion or replication • Very high potential for additional benefits (e.g. scale-up, multiplier or replication) identified	
	Theory of Change (Weighting X 3)	• Little or no likelihood outputs will deliver purpose • Too little information on assumptions to assess effects on outcomes • Risk of not achieving purpose very high	• Outputs do not meet 'necessary and sufficient' rule • Assumptions are questionable and insufficiently detailed • Risk of not achieving purpose high	• Outputs are necessary and sufficient to deliver purpose Some assumptions about externalities realistic and credible; some questions about coverage and/or depth • Some risk of underachieving but managed to enable achievement of purpose	• Outputs are necessary and sufficient to deliver purpose • Realistic and credible assumptions about externalities, good coverage and depth • Low risk of underachieving; likely will achieve purpose	• Outputs are necessary and sufficient to deliver purpose • Realistic and credible assumptions, analysing key externalities in sufficient depth • Probable will achieve or exceed purpose	
	Relevance and Robustness of Indicators (Weighting X 3)	Indicators are largely neither relevant nor robust. Relevance=clear, rule-driven, causally linked, gendered, pro-poor and cross-sectoral. Robust=data to support indicators (and base-line) are available, accessible, credible, own-able and disaggregate-able.	Indicators have many significant weaknesses in terms of relevance and robustness. Relevance=clear, rule-driven, causally linked, gendered, pro-poor and cross-sectoral. Robust=data to support indicators (and base-line) are available, accessible, credible, own-able and disaggregate-able.	Indicators have some significant weaknesses in terms of relevance and robustness. Relevance=clear, rule-driven, causally linked, gendered, pro-poor and cross-sectoral. Robust=data to support indicators (and base-line) are available, accessible, credible, own-able and disaggregate-able.	Indicators are mostly relevant and robust. Relevance=clear, rule-driven, causally linked, gendered, pro-poor and cross-sectoral. Robust=data to support indicators (and base-line) are available, accessible, credible, own-able and disaggregate-able.	Indicators are relevant and robust. Relevance=clear, rule-driven, causally linked, gendered, pro-poor and cross-sectoral. Robust=data to support indicators (including base-line) are available, accessible, credible, own-able and disaggregate-able.	
Efficiency	**Productivity measure** (Weighting X 3)	• Cost of activities/outputs higher than similar programmes & no mitigating factors identified • No evidence that value of outputs is optimised • Weak or no activity schedule and	• Cost of activities/outputs higher than similar programmes and few mitigating factors identified • Little evidence that value of outputs is optimised • Activity schedule and milestones insufficiently well planned for delivering	• Cost of activities/outputs comparable with similar programmes • Some evidence that value of some outputs is optimised (e.g. through timing of delivery, increase in proportion of output; decrease in proportion of input) • Activities planned in	• Cost of activities/outputs comparable with similar programmes • Good evidence that value of some outputs is optimised (e.g. through timing of delivery, increase in proportion of output; decrease in proportion of input) • Integration and sequencing of activities supports delivery and	• Cost of activities/outputs comparable with similar programmes • Strong evidence that value of critical outputs is optimised e.g. through timing of delivery, increase in proportion of output; decrease in proportion of input	

<table>
<tr><td></td><td></td><td>milestones
• Weak, no efficiencies and very poor input-output ratios</td><td>timely outputs
• Poor productivity, with no efficiencies achieved</td><td>integrated, sequenced way but milestones poor on timing and delivery
• Adequate productivity with some efficiencies achieved</td><td>measurement of productivity (actual ÷ planned)
• Efficient with good inputs-outputs ratio and performance likely</td><td>• Integration and sequencing of activities supports delivery and measurement of productivity (actual ÷ planned)
• Very efficient with high productivity ratio and performance expected</td><td></td></tr>
<tr><td></td><td>Risk Analysis and Mitigation
(Weighting X 3)</td><td>• Poor conflict analysis, not sufficiently conflict-sensitive
• Risk analysis poor
• Unsatisfactory monitoring tools and planning for risk mitigation
• Does not address or manage risk</td><td>• Conflict analysis weak or incomplete; some doubt about conflict sensitivity
• Risk analysis weak
• Few monitoring tools identified and some planning for risk mitigation
• Partially addresses and manages risk</td><td>• Conflict analysis captures key trajectories, drivers, patterns & power relationships that feed conflict; conflict-sensitive
• Risk analysis covers main threats to programme outputs and purpose
• Monitoring tools described; planning includes risk mitigation strategies and making timely adjustments
• Addresses risk of negative impacts and manages risk</td><td>• Grounded in well researched and comprehensive conflict analysis; conflict-sensitive
• Risk analysis covers main threats and provides good assessment of overall risk level
• Monitoring tools and planning includes risk mitigation and making timely adjustments
• Addresses risk of negative impacts and balances inherent risks with returns expected</td><td>• Grounded in well researched and comprehensive conflict analysis; conflict-sensitive
•
• Risk analysis covers key threats and provides comprehensive assessment of overall risk level. Monitoring tools and planning includes risk mitigation and making timely adjustments across identified activities, modalities and partnerships
• Addresses risk of negative impacts and balances inherent risks with returns expected in significantly difficult circumstances .</td><td></td></tr>
<tr><td>Economy</td><td>Procurement
(Weighting X 2)</td><td>• No discernable use of procurement to manage or reduce costs</td><td>• Some identifiable management of costs through procurement
• Ongoing monitoring of procurement costs not identified
• Little or no assessment of effect of procurement savings on outputs/outcomes
• Costs are managed through procurement</td><td>• Costs managed and increased economies identified through procurement
• Ongoing monitoring of procurement costs planned
• Risks to outputs/outcomes identified
• Costs are managed and reduced through procurement</td><td>• Costs reduced, and supported by evidence of savings achieved through better use of procurement
• Ongoing monitoring of procurement costs planned
• Risks to outputs/outcomes identified and assessed
• Costs are managed well and effective savings found</td><td>• Significant cost reductions achieved through better use of procurement, supported by evidence
• Ongoing monitoring of procurement costs planned
• Risks to outputs/outcomes identified, assessed and minimized
• Costs are significantly reduced and managed to very good effect</td><td></td></tr>
</table>

	Unit Costs (Weighting X 3)	• Very high cost compared with benchmarked unit cost (BM) • No mitigating factors identified which explain and justify additional cost • Cost exceeds BM by wide margin, and represents poor return	• Cost is above BM • Few mitigating factors explained which justify additional cost • Cost exceeds BM and is not delivering adequate returns	• Cost comparable with BM • No additional benefits identified • Cost is comparable and delivering adequate returns	• Cost comparable with BM • Some additional benefits described and quantified • Cost is comparable and represents good return	• Cost is below BM • Some additional benefits described and quantified • Cost is lower by wide margin and represents excellent return	
	Totals						**MAX= 100**

Testing the prototype

The table of criteria was presented through a short exercise testing its relevance and utility for programmes at the Governance and Conflict Advisers' Conference on 25th November. Points made by participants during the exercise are summarized in Appendix 3: Comments on Table of Criteria (rating and weighting). They indicate overall that the table was well received and could be easily used in a number of ways as a support guide, and as a result participants felt more informed, encouraged and confident in conducting VFM assessments. In addition, many comments indicated that the table is likely to be taken and used by participants with immediate effect for work with their teams in various 'quality assurance' activities.

Specific points to note for further work on VFM arising from this feedback are:

Links to the Business Case template and guidance need to be made. The criteria and the rating process add considerable substance to the procedure in its current form, providing a structured, systematic way to analyse programmes for VFM and to prepare a business case. A representative of FCPD welcomed the development which was seen to complement and supplement the latest draft version of the Business Case template and guidance. In particular, it addresses deficiencies in the current case, pointing to the kind of information that will be required to complete it. As noted previously, the Business Case template and guidance have been developed in parallel to this piece of work, and so to date, it has not been feasible to ensure a proper read across.

Information sources are not readily available. Several comments indicate that logframes and project memoranda have limited information to complete a business case, but some e.g. unit costs, may be found during implementation. Others noted that information is the responsibility of different staff members within the programme cycle – such as procurement. In a sense, however, the table is useful in highlighting the current deficiencies and areas which might require further work.

Utility of the table is a strong positive finding. Although there are many suggestions for improvements to the criteria, all focused on strengthening it rather than on an alternative. There is a tension between including more detail, (and specifically refining the table to include more relevance to context, political analysis and the strategic aims of governance), and keeping this generic and applicable to all DFID programmes, The introduction of the Business Case template in January 2011 is certainly focusing minds and effort on being ready for this major change - and it may be more immediately useful and motivating for staff to incorporate a generic set of criteria into the guidance at this point (see the Conclusions section for more details).

Furthermore, it was noted how the criteria could be developed in different ways, such as: (i) A more low-tech version that provides a checklist for Advisors and other staff to review management processes and delegate responsibilities (e.g. within a country office setting); or say, (ii) Adapting the table for use by an approval committee – where thresholds (or minimum requirements) would be necessary to ensure that projects do not 'pass' by scoring highly against a criteria that his heavily weighted, and yet do poorly against other criteria. In a committee setting, the table could provide a systematic means of reviewing *ex ante* appraisal documents, whereby approval might also be on a conditional basis (e.g. where minimum requirements are not met, and mechanisms should be put in place before implementation).

There are some areas that remain to be addressed to incorporate this into the Business Case. Some of the feedback indicates that the theory of change criterion needs more detailed work to address criticisms of the descriptors (too linear, too focused on the logframe). Other comments point to more work on the use of evidence in the assessment, and adding in important missing criteria such as duration of benefits. Weighting needs much more consideration, both in apportioning value to different criteria, and to ensure flexibility in the weighting process, perhaps through teams valuing together what they are trying to do.

On the road to VFM: how far does the table of criteria take us?

It is acknowledged that the prototype tested on 25 November is not a full cost benefit analysis of the kind typically found in VFM appraisal procedures. What has been proposed for rating and weighting does not attempt to replace it, or redirect VFM away from such an analysis – though in some cases a full cost benefit analysis may not be the most appropriate approach. The table provides a practical and necessary interim step for Advisers to improve VFM in programmes where possible, given what already exists in the organization in terms of procedures, information systems, management practice and staff readiness and ability, and does so in a way that is congruent with accepted programme development practice. It is based on the recognition that DFID is adjusting to new procedures, and is well placed given a good foundation of results-based management to develop an approach to VFM within which more elements can be assessed with increasing rigour.

5 Conclusions and next steps

- **Implementing the Business Case template and guidance**
- **Development of comparative measures for economy and efficiency**
- **Institutional development**

Implementing the Business Case template and guidance

Based on the work already done and presented to DFID on 25 November, the criteria that have already been developed can be further refined to support implementation of the Business Case procedure for approving programme expenditure. In this section, we highlight some of the ways forward.

A) **Strengthen the VFM assessment in the short term**. This is a relatively quick exercise; to update the 'ratings and weightings' table to incorporate key points from the feedback, keeping to a generic framework for programme assessment. This will require ensuring full alignment with the Business Case template, and some work on determining how 'rating and weighting' is used to inform decisions in the submission and assessment process as to whether to go ahead or not with the programme. An accompanying guidance on assessment methods and thresholds can be produced, to support rollout of the Business Case. This would benefit from a testing phase with a core group of staff (e.g. a DFID-wide steering group), including some from Country Offices and ideally with FCPD (given the feedback on the wider applicability of the weightings and a ratings approach).

B) **Pilot and further improve the VFM criteria.** This would strengthen the value for money case of a programme and increase the rigour of VFM assessments. This approach takes the view that the 'ratings and weightings' approach should not only be seen as a 'quick fix' to address the new Business Case/ Operational Planning requirements, but instead becomes part of a more mainstreamed 'way of doing business' – the core of which becomes part of sound management practice, and is not just a response to the current institutional demands. This will require the development of further resources to support the 'ratings and weightings' approach, covering for example more criteria, refining descriptors, use of criteria, use of indicators, sources of evidence, consideration of assumptions and risks. Ideally this would involve piloting the approach within one or two country programmes to test the reliability of the methods. A 'How To Note' could be one key output developed from further work in the medium term – either in its own right, or as part of a wider VFM How To Note, or as part of the guidance on the Business Case. This would need to coordinate with the VFM Team in FCPD.

C) **Guidance for Theories of Change for Governance and Conflict.** Further work is also 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. Without a robust and explicit theory of change, the Suggested Indicator list runs the risk of becoming isolated from meaningful interventions and their application (i.e. a sort of 'pick 'n' mix' menu approach). One way of addressing this issue is to develop guidance on writing and expressing theories of change, which would have specific value to completing the Business Case.

D) **Supplementary guidance on baskets of indicators.** At the Sunningdale meeting, there was particular enthusiasm for the idea of a basket of indicators. There is a 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.

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'). The following suggestions are recommended:

E) **Optimising economy.** There is a need for further work across Governance programmes, analyzing both those that are Governance-focused and ones which include an important component of governance (e.g. health and education), and identifying relevant, discrete cost comparison units. These are necessary for comparing how effectively programmes minimise costs.

F) **Optimising efficiency.** An obvious gap is the lack of information about which pathways to successful outcomes are the most cost effective? Governance and Conflict Advisers would benefit from developing examples of programmes that are demonstrably more cost effective in their choice of different activities and outputs to achieve similar outcomes. Simple input-output ratios of productivity are not easy to devise for Governance programmes and those in conflict-affected and fragile states. One way around this is to produce strong and robust evidence of cost effectiveness by developing rigorously designed case studies of programmes working to similar outcomes - and using these to track, evaluate and compare their performance.[13] This could be done strategically by choosing a sub-sector of Governance, and/or by developing a case study in conjunction with another sector (e.g. health or education). It could also draw on research and impact evaluation work that is being initiated in other parts of DFID (EvD, RED).

G) **A case book of VFM.** There is an idea, directly analogous to the suggested indicators, which would be to develop a portfolio of examples of managing for VFM

---

[13] By case studies, the reference here is to robustly designed studies that are able to track key VFM elements and provide findings that compare the performance of interventions working to similar outcomes – and in a way that informs Advisors on approaches that provide demonstrably better VFM.

in governance. This case book could help inform ways to do VFM in new projects, and draw on the work on optimising economy and efficiency (as described above).

Development of Options A-D

Programmes vary in their readiness to adopt different options for strengthening VFM from improving results based management through to cost-benefit analysis. They are likely to require more detailed support and guidance. It is recognised that some country offices have developed more capacity than others to move towards CBA. It would add considerable value to the work done so far by developing the more challenging options further. Suggested approaches are given below.

H) **Produce guidance on how to select the right option.** This would progress to using all the different options to demonstrate VFM at the design stage, and linking to impact evaluation studies and to other more recent developments. There would be benefit in making available some general guidance on the four options for improving value for money as described in the report, which would cover for instance:
- How to select the right option.
- How to use each option to strengthen the value for money case at the programme design stage (*ex ante*).
- How to use each option to demonstrate/assess value for money at the evaluation stage (*ex post*)
- What a VFM case might look like – how to formulate *ex-ante* and *ex-post* judgements on value for money and translate these into recommendations.

I) **Develop Option C and Option D.** This would build on the above to capture outcome benefits. Particular attention should be given to the development of Option C as a *measurement* option and as an important alternative for use when CBA/SROI are unworkable or implausible (i.e. would invite unhelpful manipulation of uncertain assumptions). This would require an initial pilot exercise to further develop a rating scale for suggested indicators, the alignment of this scale with logframe targeting procedures and a review of an existing sample of evaluation reports to demonstrate how performance might be RAG (traffic light) rated in line with the scaled indicators. Such an Option C would address the major challenges facing programmes working in conflict affected and failed or failing states, where deterioration in conditions often mitigates against progress. In such circumstances, being able to demonstrate achievement in preventing worse outcomes is critical to determining value for money. Option D would allow *some* elements/ components of *some* programmes to begin to perform cost benefit calculations of the kind envisaged in SROI guidance, where appropriate. It would also address the important question of stakeholder involvement that is essential to Governance programming, and to wider consideration of the Paris principles involved in harmonizing donor approaches, ensuring that Ministers and other key stakeholders in developing countries are maximizing opportunities to manage aid (as opposed to being recipients of aid).

Institutional development

In all the above, it will be critical to address institutional coherence. There are clearly implications for wider research, reporting, communication and capacity building across DFID. Key links need to be made among programme themes, corporate performance and evaluation. Given the raft of changes now in train in DFID, some consolidation of efforts and initiatives would assist staff by ensuring they complement and supplement one another, and hence contribute to buy-in and ownership at all levels. More specifically, these might include:

- **FCPD:** Further research is needed on developing a VFM definition and framework in cooperation with FCPD. There would need to be careful alignment with existing HM Government frameworks and guidance - and the associated adaptation and integration of the 'prototype' ratings and weightings tool for use with the Business Case and Operational Plan. This would benefit from a testing phase with a core group of staff, and FCPD may need to develop a training package and/or internet training options (or CD-ROM materials) for roll-out.
- **Programme appraisal and approval:** the 'prototype' tool could be developed for use as an assessment methodology, either formally adopted or informally used as an internal check by Country Offices. This would involve developing the criteria further, the weighting and minimum thresholds – with accompanying guidance on definitions and thresholds. It might also be suitable for use in more closely guided Annual Review and Project Completion review processes.
- **VFM has implications for evaluation** and DFID's Evaluation Department may wish to consider how VFM fits within their training plans and practical guidance – especially given their revised role supporting evaluation capacity within DFID.

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. <u>This is a list of suggestions, and should be treated as such</u>: 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 of 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. <u>This list will remain in draft for the foreseeable future</u>, 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).[14] 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.

[14] Dec 2009 Investment Committee papers on Value for Money

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 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; 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; +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.[15]

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.

[15] DFID How To Note 'Standard Indicators' (October 2009)

Appendix 2: List of reviewed documents and website materials

Andrews, M., *The Good Governance Agenda: Beyond Indicators Without Theory*, Oxford Development Studies, volume 36, 4, 379-407, 2008

African Development Bank *Results Measurement Framework 2010-2012*

Asian Development Bank, *Guidelines for the Preparation of Country Assistance Program Evaluation Reports*, Operations Evaluation Department, ADB, February 2006

Asian Development Bank, *Quality Review and Assessment Checklist for Projects*, ADB http://www.adb.org/Documents/Guidelines/guidelines-preparing-dmf/dmf-checklist.pdf accessed November 12 2010

Audit Commission, *Use of Resources Framework, Overall approach and key lines of enquiry*, Audit Commission, October 2009

Chartered Institute of Public Finance and Accountability (CIPFA) Financial Management Model

DANIDA *Value for Money Audit of Projects*, EQI Final Report 2008

Clarke, J. and Mendizabal, E., *Maximising the Impact of DFID Influencing - How Far Can We Go?* ODI and RAPID, July 2010

BIS Impact Assessment Toolkit, Better Regulation Executive, April 2010 www.bis.gov.uk/ia-toolkit

Department for Education and Skills *Evaluation and Appraisal Guidance, v2,* accessed 12 October 2010

DFID Aid Transparency Guarantee

DFID Briefing Paper A: *Analysing Conflict and Fragility, Working Effectively in Conflict Affected and Fragile Situations*, March 2010

DFID *Building Peaceful States and Societies*, Practice Paper, 2010

DFID Business Case template, draft, November 2010

DFID Business Case Annex B *Commercial case supporting guidance*, November 2010

DFID Business Case Annex C *Management case supporting guidance*, November 2010

DFID Business Case Annex D *Guidance on assessing robustness of evidence*, November 2010

DFID *Draft Structural Reform Plan*, July 27 2010

DFID *Governance Portfolio Review*, *Draft,* October 2010

DFID *Governance, Development and Democratic Politics*, 2007

DFID *How To Note: Economic Appraisals,* February 2009

DFID *How To Note*: ***Writing a Business Case***, draft Guidance, October 2010

DFID *Independent Aid Watchdog* press release, October 2010

DFID *Interim Guidance Note* ***Measuring and managing for results in fragile and conflict-affected states and situations***, 2010

DFID *Nigeria: VFM Action Plan – FY 10/11*, October 2010

DFID *PMF Objective on Value for Money and Results – Programme,* (no author, undated)

DFID *UK Multilateral Aid Review*, TORs, 2010

DFID *VFM Spot check on Economic Appraisals and Logframes*, Value for Money Dept, March 2010

DFID VfM and DFID Nigeria Programmes: *Strategy, Action Plan and VfM 'How to',* Gerry, A., April 2010

DFID White Paper (2006), *Making Governance Work for the Poor*

GSDRC Helpdesk *Report on value for money*, 24 Sept 2010

HM Treasury *The Green Book*, http://www.hm-treasury.gov.uk/greenbook

HM Treasury *Public sector business cases using the Five Case Model: a Toolkit*, Flanagan, J. & Nicholls, P., July 2007, accessed 18 October 2010

IDeA *Corporate Peer Review Benchmark*, Improvement and Development Agency, May 2009 www.Idea.gov.uk

IEG, *Cost-Benefit Analysis in World Bank Projects*, Fast Track Brief, June 2010 http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/CF83C56C2C21A14E8525779200783AF2/$file/FTB-FY2010%20CBA_desk-to-desk_06_29_10.pdf

IEG, *Performance in Government, The evolving system of performance and evaluation measurement, monitoring and management in the UK*, no. 24, World Bank, November 2010

NAO Briefing: *Work of DFID in 2009-10 and its priorities for reform*, report to House of Commons International Development Committee, November 2010

NAO *Analytical framework for assessing value for money*, 2010

NAO Report: *DFID, Operating in insecure environments*, October 2008

Natsios, A. (2010) *The Clash of the Counter-Bureaucracy and Development*, CGD, July 2010

OECD/DAC *Donor approaches to Governance Assessments: A Sourcebook,* Jensen, R.I. et al, 2009

SROI Network: *A Guide to Social Return on Investment*, April 2009

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

UK Audit Agencies, *Value For Money in Public Sector Corporate Services*, Audit Commission et al., National Audit Office, 2007

USAID *Measuring Effectiveness to Improve Effectiveness*, Briefing Note 5, June 2010 http://www.countrycompass.com/_docs/policy_briefs/Briefing_Note_5_Effectiveness.pdf

Appendix 3: Comments on Table of Criteria (rating and weighting)

(From the DFID Training Day, 25 November 2010)

Making links: incorporating criteria into the development of the Business Case template and guidance

- *The table points to the information required to 'make the business case'*
- *Table needs to be transposed in line with the Business Case template*
- *A systematic, structured way to 'do' VFM, useful for analyzing and writing the VFM case under the 3Es*
- *For weighting, this needs to follow the Business Case template*
- *Highlights deficiencies in existing documentation*
- *At what point is a decision made? When does a programme get the go ahead, or not?*

What information is required to complete assessments using the table; from where can it be found, and how?

- *Existing documents (logframes and project memoranda) do not provide relevant information*
- *Insufficient information on procurement, none on unit costs*
- *Can find information at certain points during the project cycle, e.g. unit costs*

Criteria – general comments on rating programmes

- *Descriptors worked well*
- *Scores can be increased through relatively simple changes, don't need to redesign the project to improve VFM*
- *Confident of our assessment using these criteria*

Comments on specific criteria (mainly theory of change)

- *Theory of change descriptors all within a results chain, should cover context, PBSB too*
- *Too logframe focused. Does not tell you if you are doing the right things*
- *Too linear*
- *Procurement criteria useful*
- *Needs to be adapted to reflect the context in which the programme sits e.g. political risk; also needs to include impact*

Criticisms of criteria and weighting

- *Very DFID-centred, what about Paris principles? Benefits to whom?*
- *Weighting should reflect the importance of different criteria*
- *Magnitude of change is not in the criteria, will 'low hanging fruit' get a better ranking?*
- *Shouldn't the quality of the evidence be not only a descriptor but also a criterion in its own right? Or perhaps fitted under Theory of Change, e.g. done before, did not work/done but difficult/done with proven effectiveness?*

- *Things that are missing: sustainability, PBSB, wider strategic aims in country context, evaluability, duration*

Application

- *This has merit and can be adapted. The Business Case tells me to go and do VFM, not how to do it. Will help identify information we will need for VFM*
- *More than a checklist, it structures appraisals, reviews in a systematic way*
- *Teams can use this to agree weighting, valuing what they are trying to do*
- *Can use this for allocating responsibilities to team members for parts of Business Case*
- *Very helpful tool*
- *A useful QA tool beyond VFM for theory of change and procurement*