



National Audit Office

Department for International Development:
Progress in improving performance
management

Contents

Contents.....	2
Executive Summary	3
Part 1: Progress Since 2002 NAO Report	9
Part 2: Corporate performance frameworks and initiatives.....	14
Corporate Performance Frameworks.....	14
Corporate initiatives	20
Part three: Mid-tier performance frameworks.....	24
Country programmes.....	24
Thematic areas	25
Multilaterals	27
Part Four: Programme Evaluation.....	30
Evaluation Policy and Function	30
Evaluation Outputs and Quality	34
Uses of Evaluation	36
Part Five: Project and Programme Monitoring	40
Project Indicators	40
Project scoring and performance management	43
Training and Guidance on Monitoring and Evaluation.....	45
Appendix 1: DFID’s Public Service Agreement 2003-06.....	47
Appendix 2: Assessment of the 2003-06 Red and Amber Targets	48
Appendix 3: Methodology	49
Annex A: DFID Response to this report	50

Executive Summary

1. This brief review of the Department for International Development's (DFID) performance management arrangements during 2008 is a follow-up to our 2002 VFM report on the same topic. It responds to a request from DFID's Accounting Officer to re-visit the topic periodically, which the C&AG agreed would be valuable. It is based on a desk review of main documents, interviews with DFID staff, and a survey of staff and stakeholder views about evaluation (Appendix 3). We did not undertake a full audit of DFID systems, some of which are in the process of being revised, and we concentrated on those areas of DFID activity most directly related to its performance targets. We drew on recent DFID reviews of monitoring and evaluation, and our findings square well with the results of those reviews.
2. DFID has responded well to our 2002 recommendations, improving the coverage and specification of its PSA indicators, streamlining the planning and performance monitoring arrangements, instituting better Board review procedures, enhancing the performance review of multilateral funding, and adding to the scale and independence of evaluation work. Perhaps most strikingly, all the staff we spoke to in DFID were fully aware of the Millennium Development Goals (MDGs) and associated Public Service Agreement targets (PSAs), and teams used the associated indicators to structure their debates about priorities and performance – in marked contrast to the findings from our earlier study.
3. Challenges to effective performance management, however, remain. The prime challenge is in securing sufficient reliable, timely data on poverty reduction outcomes and service delivery outputs to underpin aid targeting and performance analysis. Although DFID is the prime bilateral supporter of developing country statistical capacity, progress in securing better, more frequent poverty-related data has been slow. DFID has been unable, for example, to set a credible baseline for income poverty in 5 of its 22 PSA countries, and has inadequate trend information in a further 4. An increasing DFID focus on fragile states exacerbates this problem. There are similar gaps in output data – data which are crucial to support performance management and value for money judgements in an area where outcomes can take many years to emerge, and may not be measured even then. Current DFID work to develop standard output indicators is welcome, if limited.

4. A related issue concerns the challenge of associating inputs with outputs and outcomes. When much bilateral aid was in the form of discrete projects, DFID could associate its own inputs with project outputs and probably intermediate outcomes as well. With increasing use of programmatic aid, those links are problematic in principle, and in practice subject to serious weaknesses in developing country statistical systems. DFID has developed a set of standard indicators and is looking to strengthen economic appraisals to help inform project approval decisions and develop cost-effectiveness benchmarks. It has created a central unit charged with promoting and assessing value for money. And it has revised and strengthened its guidance on the creation of monitoring frameworks, requiring explicit baselines and a “pro-rating” of DFID’s share of benefits according to its share of inputs. The challenge will be for DFID to maintain the momentum of these initiatives with decreasing staff resources but increasing programme spend.
5. There are two further areas where, given the significance DFID attaches to them, measurement practice is not sufficiently advanced. First, measurement of the organisational capacity of development partners is insufficiently structured or scaled, given that so much DFID aid incorporates explicit capacity building elements. While there are a number of ratings and assessments that are used in specific sectors, and some overall “government effectiveness” indicators, DFID has no corporate approach to measuring capacity built or sustained. Second, it continues to struggle to measure the impact of its “influencing” work. Sorting out the critical influences on changes in policies or practices presents technical challenges. Initial DFID measurement trials took place six years ago, in the context of its work with multilateral organisations. And DFID has recently completed further pilots looking to apply approaches designed to measure projects. An evaluation of these pilots confirmed the value of DFID’s ‘logical framework’ approach, while noting some limitations. Performance measurement in this area is hindered by lack of good measurement of inputs: DFID has no staff time booking system that allows it to monitor the scale, cost or nature of resources committed to “influencing”.
6. DFID has well-established systems for specifying aid objectives, and associated indicators, in a logical hierarchy, and for using those indicators in approving and monitoring its programmes. But its management of these systems has not been sufficiently firm. A number of our recent audit studies have commented on problems with PRISM, the former IT system used to store and collate project and programme data, as well as weaknesses in specific project or programme objectives. In this review the projects we examined had many different indicators of progress, often leading to subjective judgements of project progress. System incentives militate in favour of more, rather than less, favourable judgements. Recently, a DFID-commissioned monitoring and evaluation (M&E) audit found widespread problems

with defective logical frameworks and weaknesses in management review. DFID has fixed some of the structural issues and is implementing a new integrated financial and project management system (ARIES). But lack of staff skills in M&E and weak in-the-line review remain real issues - staff told us, for example, of a lack of training in M&E. And they also noted that potential out-of-the-line review bodies, such as the new Investment Committee, would need to address portfolio review and systemic issues by reference to a sample of actual programme proposals, if they were to help secure better compliance with DFID policies on programme appraisal and monitoring.

7. DFID's Evaluation Department (EvD) has increased in scale and stature since 2002, and the recent creation of an Independent Advisory Committee on Development Impact (IACDI) provides a further buttress to evaluation independence and quality. EvD remains less structurally independent than many other development agency evaluation units, however, and its recent programme has been slanted towards practice and strategy reviews, rather than focusing on DFID impact, or securing better information on aid cost-effectiveness. Work is in hand to refresh evaluation policy and strategy. Staff and stakeholder perceptions of DFID evaluation are favourable, albeit with clear scope for improvement. Evaluation is not, however, well integrated into general DFID performance management – it is not, for example, tasked with providing a better interpretation of DFID's performance against PSAs or MDGs, and rarely set to work on cost-effectiveness topics. External stakeholders ranked evaluation's potential contribution to accountability more highly than did DFID staff.
8. In an area where outcomes often take many years to emerge, can be difficult to measure, and are subject to many influences other than DFID, programme implementation (and M&E in particular) has attracted less kudos than policy analysis, programme planning and fire fighting. DFID's 2008 Results Action Plan is designed to address this situation, and is a high corporate priority which will be closely monitored by top management. Stronger management arrangements for performance review have now been put in place to drive the department-wide focus on results with more formal divisional performance review and challenge arrangements, as opposed to structuring scrutiny (and therefore perceptions of credit or recognition) around policy or planning papers.

Overall Conclusion and Recommendations

9. DFID has successfully restructured its performance management arrangements to make sure there is a coherent performance debate around the adoption of aid policies and practices and the commitment of resources. It has improved general scrutiny of progress in reducing poverty, and of progress towards divisional goals. There is a lot of ongoing staff work to improve performance management, but still clear scope for further improvement. Partly this circumstance reflects a very broad DFID remit, which makes it inappropriate to focus performance assessment in the way that is possible with more limited priorities. To build on recent progress DFID should:

On data and indicators

- Give added priority to its work to improve the data available on results, and clarify its policy on the generation of data when national sources cannot provide them. Without better data it is difficult to see how performance management based around the achievement of quantified goals can be further improved or streamlined.
- Improve the ability of costing systems to link resources to performance objectives and to results. And devise simple rules to estimate DFID's contribution to improved poverty outcomes, according to DFID inputs. While DFID aid policies reflect explicit consideration of cost-effectiveness, there is also a need to obtain evidence of value for money achieved in their implementation.
- Devise a structured approach to measurement of capacity, to provide a common starting point for the various aspects of capacity-building in developing countries that DFID often supports. And look to define an approach to measurement of influencing, and consider the introduction of time-booking for its staff.
- Improve its use of comparisons, within and between programme areas, and with other development agencies. One of the key uses for measurement should be in identifying or confirming opportunities to improve performance. DFID should give special emphasis to the quality of measurement in relatively low-spend, potentially high-impact areas, such as trade; and also around innovation – to establish whether it does indeed “work”.

On performance frameworks

- Strengthen the “balanced scorecard” format for DFID that covers results and capability dimensions, by improving the understanding of the links between the various elements, and by promoting interpretation of the scorecard as a whole. Review the organisational performance frameworks adopted or implied by various assessment approaches across DFID

(including DFID's own Divisions, multilaterals, major NGOs, and other development partners), looking to standardise the frameworks where possible.

- Build on its work to assess the performance of multilaterals, and improve the management of DFID's multilateral funding, by continuing the development of donor-led systems to assess multilateral performance and capacity, and by clarifying DFID expectations of multilaterals' own performance measurement systems.
- Further revise guidance on project and programme monitoring frameworks to make clear that the logic model needs to be articulated, and to clarify which output and purpose indicators represent the basis for approving the investment - and on which ultimate performance will be judged - and which indicators are merely supporting metrics or represent "nice-to-have" elements.
- In view of the poor state of monitoring frameworks established in DFID's M&E review, improve the M&E training for all relevant DFID staff, and secure appropriately independent technical review of the frameworks for proposed project and programme - proportionate to the scale of the proposed investment.
- Consider the merits of independently validating project scores, and whether the strength of the scoring system is adequate to support the stress placed upon it by targets for improved average scores.

On evaluation

- Continue to improve the effective independence of the evaluation function, and agree with IACDI a policy for evaluation within DFID, and for resourcing evaluation work. Make sure DFID staff understand the implications of the policy, and that it is embedded in DFID management requirements.
- Improve staff awareness of evaluation reports – particularly reports such as synthesis reports, designed to have relevance to a wide range of staff. And look to secure greater prominence for summarised findings and evidence, as a complement to recent efforts to track the response to evaluation recommendations.
- Look to increase the proportion of evaluations which provide evidence on the cost-effectiveness of aid activity, or are directly relevant to key management decisions on the balance of programmes, or the choices between aid instruments or approaches. Given the large volume of evaluative material produced globally on development issues, not all of it of high quality, fewer, deeper or more decision-targeted evaluations are most likely to add value.

- Give added priority to arrangements which facilitate more joint or multilateral approaches to evaluation, where that will secure high quality evaluation and reduce net burdens.

10. DFID's initial response to this report is at Annex A.

Part 1: Progress Since 2002 NAO Report

Context: Performance against Targets

- 1.1 The international community articulated in 2000 the aim of eliminating world poverty in the Millennium Development Goals (MDGs). The overriding objective is to reduce by one-half the proportion of people living in extreme poverty by 2015 through achieving universal primary education, promoting gender equality, reducing child mortality, improving maternal health, combating HIV/AIDS, malaria and other diseases, ensuring environmental stability and developing a partnership for global development. The United Kingdom's contribution to this objective is led by DFID.
- 1.2 DFID strategic performance is assessed by reference to Public Service Agreement (PSA) Targets and accompanying measures. The PSAs are broadly aligned with the MDGs and are reviewed every two to three years, as part of UK Government's Spending Reviews. Progress towards the PSAs is reported by DFID twice yearly - in the Annual Report published in April/May and the Autumn Performance Report published in October/November.
- 1.3 The results against the 2003-06 PSA targets are at Appendices 1 and 2, and the latest position on the 2005-08 targets is shown in Figure 1 below.

Traffic Light System:

Green: Met/Ahead/On course

Progress on the target is either exceeding or in line with plans and expectation or:

The target has been achieved by the target date if we are providing a final assessment.

Amber: Partly met/Broadly on course – Minor slippage

Progress on the target is broadly on course but that there has been minor slippage. Alternatively progress may have been made in some areas but not in others or:

Target has been partially met, i.e. some, but not all, elements have been achieved by the target date, or were very close to achieving the target.

Red: Not met/Not on course – Major slippage

Progress on the target is not on course and there has been major slippage or:

Target was not met, or was met late.

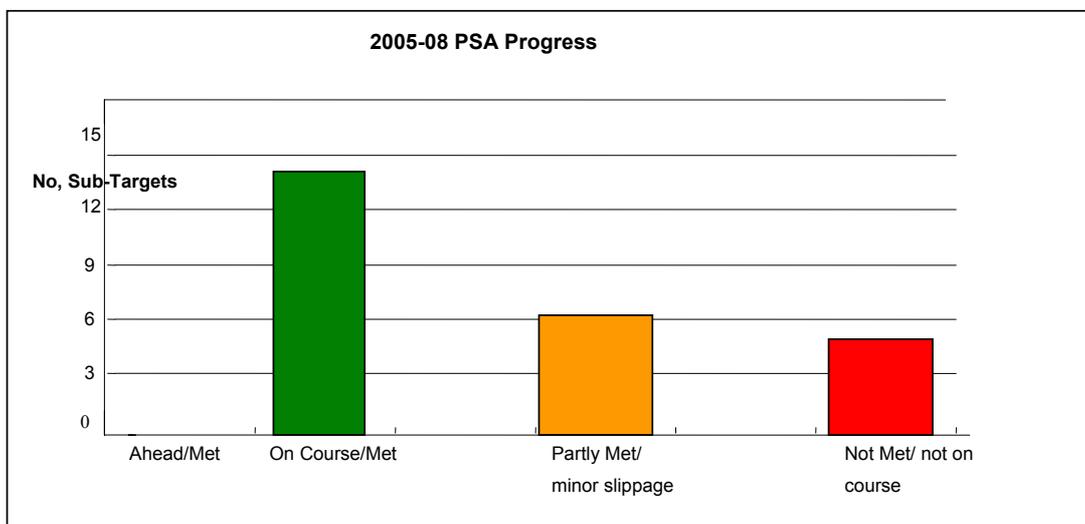
Figure 1: DFID's Public Service Agreement 2005-08

Aim:	<i>Eliminate poverty in poorer countries in particular through achievement by 2015 of the Millennium Development Goals</i>	
Objective I	Reduce poverty in sub-Saharan Africa	
Target 1	Tackle poverty in Africa <i>Progress towards the MDGs in 16 key countries</i>	Amber
Objective II	Reduce poverty in Asia	
Target 2	Address poverty in Asia <i>Progress towards the MDGs in 9 key countries in Asia</i>	Amber
Objective III	Reduce poverty in Europe, Central Asia, Latin America, the Caribbean, the Middle East and North Africa	
Objective IV	Increase the impact of the international system in reducing poverty, preventing conflict and responding effectively to conflict and humanitarian crises	
Target 3	Make multilaterals better at delivering aid <i>Improved effectiveness of the international system</i>	Amber
Target 4	Help poor countries benefit from international trade <i>Ensure that the EU secures significant reductions in EU and world wide trade barriers by 2008 leading to improved opportunities for developing countries and a more competitive Europe (joint target with BERR (Department of Business, Enterprise and Regulatory Reform))</i>	Amber
Target 5	Conflict prevention and post-conflict resolution <i>By 2007/08, improved effectiveness of UK and international support for conflict prevention, through addressing long-term structural causes of conflict, managing regional and national tensions and violence and supporting post-conflict reconstruction where the UK can make a significant contribution, in particular Africa, Asia, the Balkans and the Middle East (joint target with FCO and Ministry of Defence (MoD))</i>	Amber
Objective V	Develop, support and promote policy that assists poverty reduction and the achievement of the MDGs.	
Objective VI	Improve the impact and effectiveness of DFID's bilateral programme	
Target 6	The 90:10 target and portfolio quality <i>Ensure that the proportion of DFID's bilateral programme going to low income countries is at least 90% and achieve a sustained increase in the index of DFID's bilateral projects evaluated successful</i>	Green

Source: Development: Making It Happen 2008 Annual Report, DFID

1.4 When examining the sub targets, which comprise specific MDG-related indicators, the picture remains mixed.

Figure 2:



Source: DFID Annual Report 2008

Figure 3: Assessment of the 2005-08 Amber Targets

Target	Assessment	Explanation
One: progress towards the MDGs in 16 African countries	Amber	Africa continues to demonstrate growth but an uneven pattern of achievement, even within sectors. Maternal mortality is off-track, for example, while reduction in under 5 mortality is ahead of schedule.
Two: progress towards the MDGs in 9 Asian countries	Amber	Asia shows progress in most areas. The percentage of people living in poverty has reduced to 12% in East Asia and the Pacific and 33% in South Asia. Progress against the target to reduce HIV prevalence rates, however, is difficult to assess as the available data has poor coverage and is unlikely to represent vulnerable groups.
Three: improved effectiveness of international system	Amber	Despite strong lobbying from the UK, insufficient support from other member states meant that sub-target one, to increase the proportion of EC ODA to Low Income Countries was missed. Sub-target three, to improve the way international partners are supporting poor countries to reach the MDGs has been met and sub-target two which aims to ensure that 90% of HIPC receive debt relief by the end of 2008 is on course to be achieved.
Four: reduced trade barriers	Amber	The main lever for meeting this target is the Doha Development Agenda. Following the failure of the Cancun trade summit, Doha negotiations restarted during 2007 and revised draft texts of agreement were published in January 2008, but progress again stalled. Progress has, however, been made in areas such as working towards a more simple, transparent and development-friendly Rules of Origin and a review of the Generalised System of Preference to improve developing countries market access to Europe.
Five: improved	Amber	This PSA has been described as partially met as two of the twelve sub-targets have

effectiveness of conflict prevention		been achieved, nine partly achieved and one not met. An area of particular success was Sierra Leone, which now has a democratically elected, stable government. The Middle East Peace Process target was not met due to Israeli military action in Hamas controlled Gaza.
--------------------------------------	--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Source: *DFID Development Making It Happen 2008 Annual Report*

- 1.5 Overall, there has been a considerable reduction in the number of people living in extreme poverty. The charts above show, however, that DFID performance against targets has been mixed, as has global performance generally against the MDGs. Broadly, Africa has fared worse than Asia, and aspects of multilateral and trade reform have also proved problematic. The global economic environment has been favourable since the early 2000s, with considerable economic growth in many developing countries. Interpreting DFID “performance” from these numbers is extremely difficult, as is the difficulty of separating the influence of aid generally on poverty reduction from other factors. We follow up this general issue in subsequent parts.
- 1.6 We looked to see whether DFID had implemented the recommendations in our 2002 report: Performance Management – Helping to Reduce World Poverty. We found that it had addressed, through responses at the time and other changes made subsequently, many of the issues we raised. The following paragraphs summarise some of the key changes made under the main recommendation headings used in our 2002 report. These changes are considered in more detail in the remaining Parts of this report.

Strategic Performance Framework

- 1.7 DFID has maintained a focus on internationally-agreed poverty indicators in its PSA agreements, using the Millennium Development Goals (which superseded the International Development Targets, but drawing on similar indicators) as their basis. Over time, the PSA has been broadened in scope to cover a greater proportion of DFID’s bilateral expenditure. And the associated indicators also cover multilateral and non-expenditure areas, if in a less quantified way.
- 1.8 Corporate objectives have been cascaded throughout the organisation. Staff have a much higher awareness of MDGs, PSAs and now Departmental Strategic Objectives (DSOs) than in 2002, with a greater understanding of their own contributions towards the achievement of corporate objectives. The MDGs in particular frame much of the debate on progress towards poverty reduction, and use of MDG-related indicators is common place in most analyses of poverty reduction performance.

Country Programme Management

- 1.9 Country assistance planning has been strengthened through the adoption of an explicit focus on poverty reduction outcomes, higher-level scenario and contingency planning and better assessment of context and risk (for example, through Country Governance Analyses).
- 1.10 Measurement of “influencing” remains in its infancy. Earlier attempts to look at influencing through multilateral and evaluation groups did not yield a common approach. Finance and Corporate Performance Division are now piloting a 'projectised' approach. Training on 'outcome mapping', one step in that approach, is being made available across DFID.

Performance Management and Review

- 1.11 DFID has improved its management reporting to the Board with a combination of results-based reporting, together with Divisional reviews of progress with planned activities, including policy reviews. In addition, there are other periodical or ad hoc exercises, such as staff and stakeholder surveys, Capability Reviews, OECD Peer Reviews, and procurement reviews which together provide a sort of “balanced scorecard” overview. New arrangements will make the links from PSA through DSOs to Divisional and lower level plans clearer.
- 1.12 Country Programme Evaluations were introduced in 2004 and cover at least 5 PSA countries each year. Findings are synthesised on an annual basis and presented to the Management Board and Country Peer Review Committee.
- 1.13 The 2007 publication Making Statistics Work for Development sets out DFID’s statistics strategy and how the measurement of results can contribute to better development policies. DFID currently supports building statistical capacity initiatives in over 20 partner countries. DFID's Evaluation Department is currently leading a multi-donor evaluation of support to this area.

Part 2: Corporate performance frameworks and initiatives

2.1 Despite the improvements made since we reported on DFID's performance management in 2002, there are ongoing challenges in respect to the corporate, middle and micro level of the organisation.

Corporate Performance Frameworks

2.2 Following CSR07, DFID has revised its performance planning and monitoring arrangements to cater for the new cross-government PSA arrangements and the introduction of Departmental Strategic Objectives (DSOs). PSA 29, on which DFID leads but pursues jointly with HMT, FCO and DEFRA, remains based around the pursuit of the MDGs. It is measured by reference to 8 indicators, on a global and "focus country" basis. DFID's DSOs look to address factors that will permit an effective DFID contribution to PSA achievement. Figure 4 below set out the main elements of the corporate measurement framework.

Figure 4

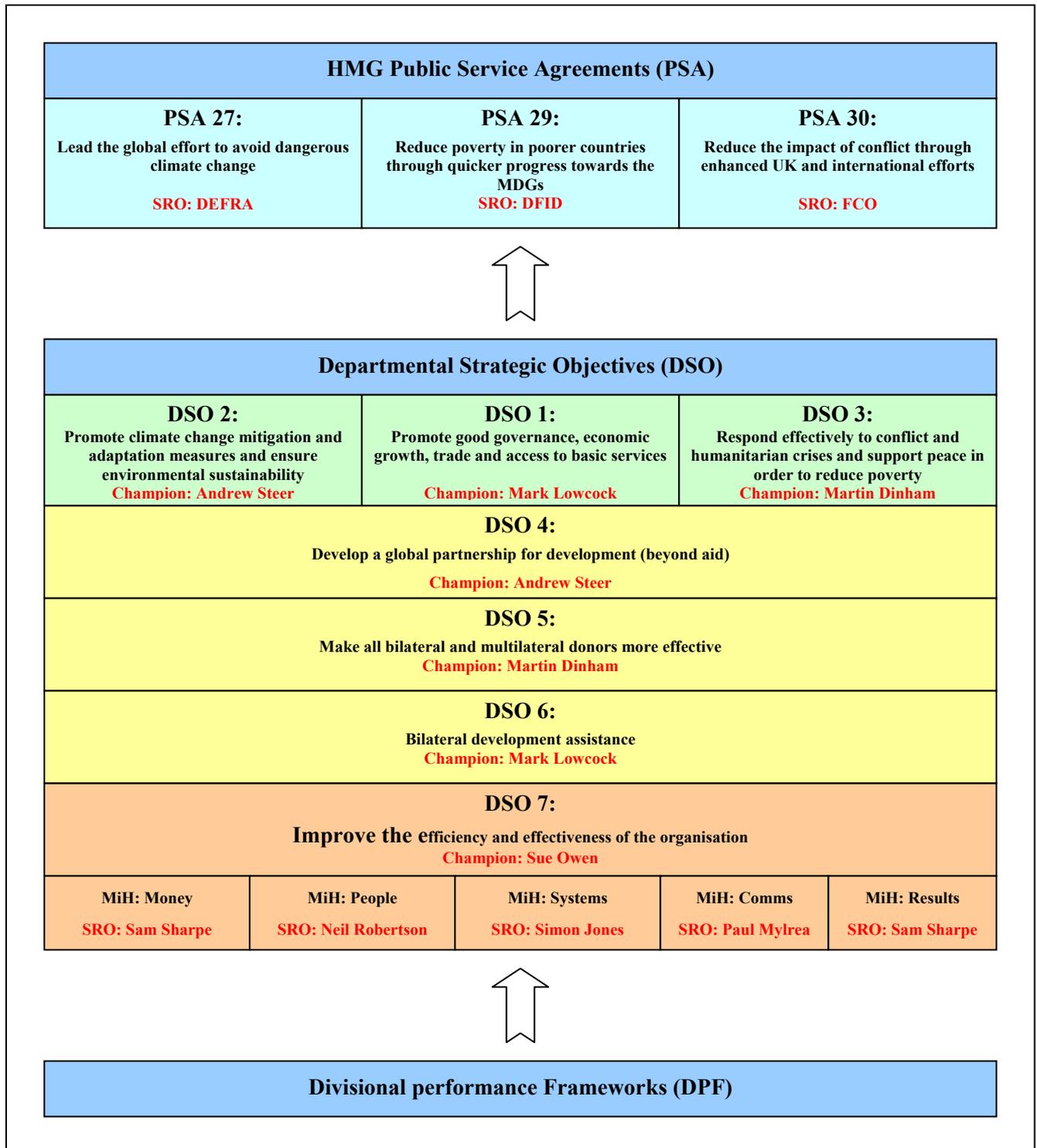


Figure 5

PSA DELIVERY CHANNELS		KEY PSA ACTIONS	RESPONSIBILITY
Working through the bilateral development programme		1. Scale up aid and other assistance to Africa and Asia in line with UK commitments while maintaining aid programme quality	DFID
		2. Increase the provision, access to and quality of essential services	DFID
		3. Support putting economic growth at the heart of developing countries' own poverty reduction strategies	DFID
		4. Help African and Asian governments and institutions achieve sustained levels of economic growth, promote trade and improve the investment climate	DFID
		5. Overcome persistent poverty among the socially and economically excluded and provide social protection for the most vulnerable groups	DFID
		6. Help improve governance	DFID
Influencing the international system	Economic Growth	7. Growth diagnostics undertaken in all countries in which DFID operates to understand bottlenecks, constraints to growth and focus assistance in those areas	DFID, HMT
		8. Press the IFIs to provide developing countries with improved, consistent and balanced policy advice on achieving inclusive growth and employment	DFID, HMT
		9. Work with business to maximise the contribution of the private sector to investment, growth and poverty reduction	DFID, HMT
		10. Work to ensure that adequate funding is available through the EC, World bank, African development Bank and other Regional Development Banks for financing infrastructure	DFID, HMT
		11. Help to develop better policy tools to improve the investment climate, including tools for business regulation, competition policy and tax and customs reform	DFID, HMT
		12. Help to promote financial inclusion and tackle barriers that prevent poor people from gaining access to markets	DFID, HMT

- 2.3 At PSA level, one indicator has been selected under each of the MDGs to act as a basis for monitoring in the 22 countries where most DFID effort is concentrated. In addition, DFID has agreed with its Departmental partners a set of 81 “key actions” in support of the PSA, and will use 16 of those actions in monitoring to assess the extent of cross-government collaboration (examples in Figure 5). PSA monitoring involves a traffic light system for each selected PSA indicator, by country, according to whether DFID judge the country to be on track to meet the MDGs for that indicator (green), off-track but with possible MDG achievement by 2040 (amber), or severely off-track (red). These assessments are used to come to a simple average judgement of status for the set of 22 countries. Because data availability is a problem in most of these countries, DFID has also created a “direction of travel” assessment, which it makes to supplement indicator-related judgement. DFID targets are of the form “maintain the status of countries judged “green” in the baseline assessment, and accelerate progress in at least x of the remaining countries”.
- 2.4 For DSOs, DFID has defined 32 performance indicators to track progress at the corporate level (examples at figures 6 and 7 below). These indicators and associated success criteria have been used to develop Divisional Performance Frameworks (DPFs) for each of DFID’s 11 Divisions (each Division monitoring progress against its own sub-set of those indicators). Progress against the DSOs and DPFs is reviewed formally every six months. For each indicator DFID provides a summary traffic-light measure of progress along the following lines:
- Red: No progress or regression
 - Amber: Some progress on baseline but not enough to meet success criteria
 - Green: Progress is consistent with meeting final success criteria
- 2.5 The traffic light assessment is made centrally by the Corporate Performance team supported by a network of DSO monitors helping to secure appropriate judgements. This assessment is then translated into an “improvement” or little or no improvement rating in line for external reporting purposes in line with Treasury guidance.

Figure 6 DSO 1.4a and b: Delivery of the White Paper commitments on public services (improved outcomes for education and health)

	Indicator 1.4a	Indicator 1.4b
Description	Primary completion rate in PSA countries	percentage of births attended by a skilled professional in PSA countries
Data provider	DfID	DfID
Data set used	DfID Country level Assessment of 100% PCR by 2021	DfID Country Level MDG Assessments
Link to external data source (if applicable)		-
Baseline data value	10 countries off track	6 countries off track
Reference period for baseline	2007	2007
Target data value	Half the number of off track countries	Half the number of off track countries
Reference period for target	2010	2010

Figure 7 DSO 5.3: Improved effectiveness of the EC

	Indicator 5.3a	Indicator 5.3b
Description	Commitments for general budget support and sector policy support programmes from the EDF	i) 40% of field missions are joint by 2011 ii) 66% of country analytic work, including diagnostic reviews, are joint by 2011; iii) 100% of donor capacity-development support provided through co-ordinated programmes consistent with partners' national development strategies, by 2011
Data provider	European Commission	European Commission

Data set used	Europe Aid Co-operation Office Indicators The indicator follows the increase in the percentage of total annual commitments earmarked for general budget support and sector policy support programmes (SPSPs) and not earmarked for classical project support. The percentage for SPSPs details the percentage for sector budget support within the SPSP	Statistical tables from "Survey on Monitoring the Paris Declaration"
Baseline data value	30.2%	Baseline figures: 33%, 48%, 33% respectively
Reference period for baseline	2006	2006
Target data value	40%	i) 40% ii) 66% iii) 100%
Reference period for target	2011	2011

2.6 These arrangements represent improvement in a number of areas, in that they:

- provide both a broad overview and a focus on countries where DFID has a bigger influence over outcomes;
- recognise weaknesses in the availability of outcome data;
- provide a clear, agreed framework for joint action;
- secure a degree of consistency between DFID planning and monitoring at different organisational levels;
- offer a basis for aggregating performance up to a corporate level.

2.7 They also suffer from a number of weaknesses:

- it remains problematic to identify DFID resources applied to specific performance objectives or indicators making it more difficult for performance to inform key resource allocation decisions;

- data for the outcome indicators remain heavily dependent on partner country systems, and underlying problems in data quality and availability limit the usefulness of the indicators in performance management;
- the selection of PSA key actions, and choice of DSO indicators, raise questions over the underlying “logic model” linking DFID activities to development outcomes.

2.8 Board reporting of performance under the new arrangements has only just started. We reviewed performance reporting arrangements under the previous system. That involved quarterly performance against PSA and other corporate targets, as well as periodic reviews of Divisional performance, and other performance-related issues, such as Capability Review follow-up and staff survey results. It was clear that the Board devoted considerable time and effort to performance review, and that there was a good degree of coherence between formal targets and the real interests of the Board. It was equally clear that while monitoring systems provided general pointers to areas of success or failure their capacity to provide more detailed analysis of performance issues was limited – there was limited capacity to “drill-down”. This latter element was more in evidence in the Divisional performance reviews, where central staff had prepared more detailed briefs for the Board to use as a basis of questioning.

2.9 Areas where the system struggles to provide relevant data include:

- DFID’s contribution to performance as part of a wider donor community. It would be helpful to have some information on overall donor inputs, and possibly activity, to help position DFID within the contribution of aid generally on the outcomes being monitored; and
- The state of the economy in the focus countries, and its contribution to the monitored indicators.

Corporate initiatives

2.10 DFID have adopted a Results Action Plan, in response to a Capability Review in 2007. The Plan promotes better evidenced based decision making and more effective use of performance information to achieve maximum impact in development effectiveness. The Results Action Plan has been given further impetus as one of five workstreams in DFID’s “Making it Happen” business change programme launched this year. It requires ten priority actions

Area	Priority Action
Within DFID	More use of quantitative data to improve decision making
	Strengthened performance and results frameworks for country programmes
	Improved communication with the public on results of development assistance
	Review of people management systems to encourage stronger focus on outcome
	Establish the Independent Advisory Committee on Development Impact to strengthen the independence of evaluation
With partner countries	Investment in statistics
	Support accountability mechanisms to scrutinise government and donor performance
Internationally	Support international approaches to impact evaluation
	Promote mechanisms for mutual accountability
	Promote mechanisms for assessing aid effectiveness

Source: DFID: Results Action Plan 2007

2.11 The creation of the Investment Committee demonstrates DFID commitment to ensure that its programmes represent good value for money. The Investment Committee is a sub committee of the Management Board. It examines strategic investment decisions and resource allocation, the balance of the investment portfolio and ensures the adequacy and implementation of investment appraisal systems to take strategic financial decisions on the basis of evidence.

Investment Committee: Terms of Reference
<ul style="list-style-type: none"> • Advise the Management Board / Ministers on strategic investment choices and resource allocation priorities, ensuring that the mix of investments is aligned with DFID's strategic vision • Provide assurance to the Management Board / Ministers that DFID programmes represent good value for money • Ensure DFID's investment appraisal procedures and practice are fit for purpose and champion the use of evidence and quantification of outputs.
<p>Source: DFID: Investment Committee Terms of Reference 2007</p>

2.12 The Making it Happen change programme was introduced to promote delivery, focus and impact on development, and to enable improvements in the way the organisation works to ensure the greatest impact on the Millennium Development Goals. There are five workstreams: Results, Money, People, Communications and Systems. The Results workstream aims to embed the issues raised in the Results Action Plan in the routine of DFID business.

Making It Happen Objectives

Delivery: leveraging development impact by working across Government, making international institutions more effective and accountable and grounding work in the direct delivery of the best possible bilateral programmes.

Focus: on the actions that will have the greatest impact on poverty including tackling growth, conflict, and climate change and on front-line work to transform priority fragile states.

Impact on Development: improving measuring impact and communicating successes to the public to sustain support for development in the future.

Source: DFID

- 2.13 DFID acknowledge that it is important to link performance and financial data to inform decisions. Indeed, at the DFID Senior Civil Servants Away Day in March 2008 there was a mandatory session on using numbers in decision making. There are two corporate projects supporting this theme. One aims to create a standard set of output and outcome indicators, the other to ensure project approval decisions are informed by more rigorous economic appraisal and more quantification of expected outputs and outcomes. Together, these changes should facilitate the development of cost-effectiveness indicators, from which some benchmarks about cost-effective performance might be constructed.
- 2.14 The standard indicators aim to secure information in a common format for programme areas, such as health, education and infrastructure, where DFID often works and where there is already a degree of standardisation through professional or multilateral influences.

Examples of Standardised Indicators

Health

- Number of hospitals built or rehabilitated
- Number of insecticide treated bed nets distributed

Education

- Newly enrolled primary school students
- Number of textbooks purchased and distributed

Water and Sanitation

- Number of people with access to new/improved water supply
- Construction and rehabilitation of drainage

Source: DFID

- 2.15 Clearly, such standardisation would have more value if adopted internationally. DFID considered asking the UN to take the project forward, but concluded that it would be difficult to achieve international consensus quickly. They hope, however, to be able to work with the EC and member countries. Another issue is that developing countries currently collect data in a variety of different formats. DFID policy is to support national statistical systems, and not to dictate statistical priorities to developing nations, or circumvent their systems by creating parallel systems. Even where DFID has programmes relevant to the standard indicators, there is no guarantee that the required data will be available. Any improvements in data coverage are likely to be slow. DFID estimates that the standard indicators might be relevant to around 40 per cent of its bilateral spend. But the patchy availability of data even within this proportion is one of the reasons DFID will not use the data for comparisons across countries.
- 2.16 The cost-effectiveness project faces the same limitations, inasmuch as it hopes to draw on the standard output indicators. It aims to secure better quality of log-frames at project and programme approval stages, with more quantification of intended inputs, outputs and outcomes relative to defined baselines. This information can then be used as a basis for cost-benefit or cost-effectiveness analysis. A recent DFID M&E review, however, noted that existing log-frames often lacked baselines or inaccurately assigned indicators to “output” and “outcome” categories. Managing the desired improvement “in the line” will present real challenges. The role of independent scrutiny of investment proposals could be a key force for change: the new Investment Committee could contribute to extra rigour, provided it scrutinises a sample of actual log-frames as part of its portfolio and systems review functions.

Part three: Mid-tier performance frameworks

3.1 In this section, we look at performance management in operational areas of DFID – in country programmes, which form the bulk of bilateral expenditure; thematic areas, where DFID impact on poverty reduction can be disproportionate to resources committed; and funding to multilaterals, which now represents over 40 per cent of DFID budget and has been increased under the CSR07 settlement.

Country programmes

3.2 DFID has made recent changes to country programme planning and monitoring to bring them into line with the new corporate performance frameworks. Country offices now set out an Issues and Choices submission for Ministers, which deals with the context and possible strategies for the programme. Once Ministers have signalled their views on the strategy, country teams must prepare a Business Plan, which includes a number of standard measurement grids, including: objective, outcome, output and activity measures; contributions to DSOs; contributions to MDGs and PSAs; and risks to performance objectives. These standard annexes are designed, inter alia, to facilitate the aggregation of performance through Divisional to corporate levels, and to minimise the need for ad hoc requests from the centre for performance data.

3.3 Previously, both strategy and various measurement requirements featured in Country Assistance Plans (CAPs). In practice, however, CAPs struggled to maintain life after the initial approval stage. DFID's M&E review found that the CAP guidance to be so vague that the Plans did not function as performance frameworks. And we found that the annual CAP reviews required in DFID's standard procedures were rarely completed.

3.4 On our visits, we found that DFID country offices took different approaches to monitoring their performance. For example, DFID Nigeria uses Quarterly Strategic Reviews, based around the four pillars of their CAP. They had recently introduced a specific challenge function into this process, under which staff were assigned roles to question the assessments in the report. While the link between DFID project and programme performance and overall Nigerian progress was highly qualitative, the DFID Nigeria team found this process helpful to identify areas requiring action, and projects or programmes peripheral to the team's main objectives. DFID Sierra Leone

use an Internal Management Framework based on a Joint (with the EC) Country Strategic Results Framework which states key objectives and indicators.

- 3.5 New guidance on country programme planning requires the production of a “results framework”, which will then form the basis of country team performance reporting and monitoring. The framework requires the identification of key performance objectives, and associated outcome and output indicators, as well as associated DFID resources. It also requires links to be made with DSOs and other DFID priorities. A project scoring summary is optional. This initiative offers the prospect of more coherent, streamlined reporting within DFID, with greater scope for comparison between country teams.
- 3.6 There remain, however, a number of challenges and tensions in assessing country team performance. Country team performance reporting – by reference to time, cost and quality objectives in projects and programmes – will centre on main themes identified by the country team. Information on smaller projects (by financial value) and those outside the main themes will not form part of the team reports. Developing nation or other partners’ views of the DFID team are not captured systematically. There is no routine monitoring of the quality of programme judgements – on which evidence would emerge only over the medium term – or of the degree of conformity with DFID policies, strategies and mandatory practices. Periodic internal audits and country programme evaluations provide insight into some of these issues, but these reviews are not frequent: for example, although 27 country evaluations have been completed since our 2001 study, the average time between evaluations in any country will exceed 5 years. There is little comparison between DFID country offices, and no ready way of rating the performance of business units now often employing 40 or more people, and with annual budgets of around £100 million.

Thematic areas

- 3.7 There is a growing recognition within DFID of the importance of policy influencing and capacity building in areas such as governance, trade and gender equality. This work consumes relatively low resources, with potentially high impact. However these areas present monitoring and evaluation challenges because of limited quantifiable data, an absence of baselines, problems in defining “logic models” linking activities to outcomes and difficulties in attributing any progress to DFID actions.
- 3.8 While thematic areas often have associated medium term strategic plans, these plans have rarely set out performance measurement requirements. As a result, monitoring and evaluation activity has often been ad hoc. DFID has recently strengthened its policy

management systems through a requirement for all policies to include implementation and review plans.

Monitoring of thematic based activities

‘We provide information for the corporate performance framework and annual report.’

‘At the strategic level monitoring and evaluation is reactive as opposed to following formal monitoring and evaluation procedures.’

Source: NAO interviews with staff in thematic policy teams

- 3.9 Under the new DSO, there are specific “indicators” for governance, trade and gender, for example, under which programme, support and policy divisions may list further targets and indicators. The nature of these lower level indicators varies widely, from activity-based indicators subjectively assessed through to quantitative outcome indicators. The line of sight, however, from policy influencing work carried out in multilateral and policy divisions to development outcomes is unclear. In some cases lower level indicators duplicate PSA indicators and the selected outcome indicators under DSO 1. In other cases, there appear to be gaps in lower level frameworks where activity is necessary to deliver against DSO indicators. In particular, the specified indicators do not consistently define or chart the key DFID activity or outputs needed to lead to the desired outcomes. Thus securing a clear view of DFID performance, as opposed to general development progress, remains problematic.
- 3.10 Given the nature of these areas, periodic evaluation studies could be a particularly useful source of management information, since monitoring indicators could prove difficult to interpret. We reviewed a sample of relevant studies to see how well they linked to performance management and policy review. We found that such links had been mixed.

Thematic Area	Evaluation	Link to policy formation
Gender	Evaluation of DFID’s policy and practice in support of gender equality and women’s empowerment (Series of thematic studies and country case studies) (2006)	The main themes of the evaluation feature in the Gender Equality Action Plan (GEAP), such as the creation of the role of multiple champions.

Trade	Changes in strategic influence: DFID's contribution to trade policy (2003)	Contributed to the Trade and investment white paper: making globalisation a force for good 2004
Trade	Evaluation of DFID's trade related capacity building synthesis report (2005)	None obvious. Trade activity was being scaled up in 2006/07 but this was not joined up with the evaluation

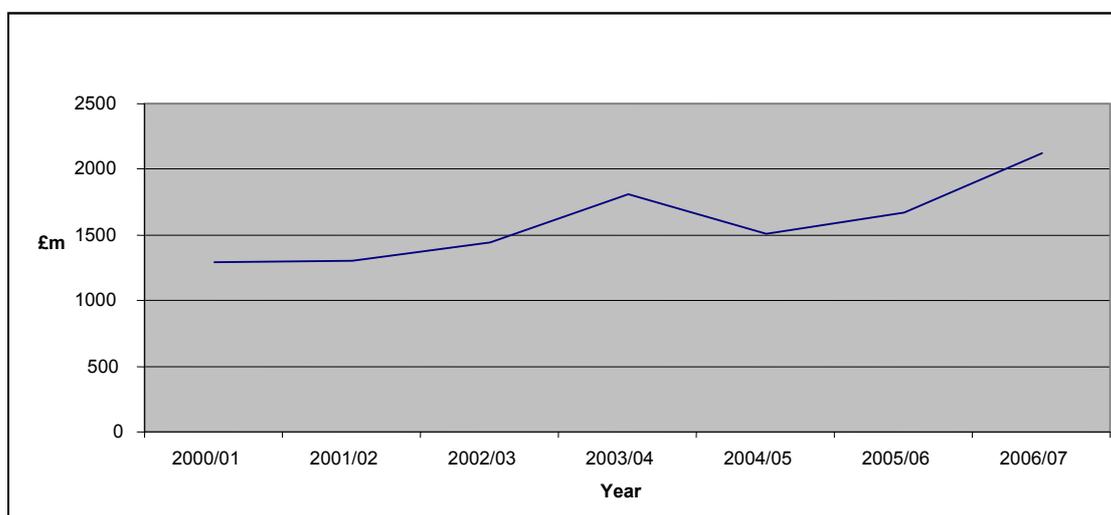
Source: NAO review of thematic based evaluation documentation

3.11 DFID's Strategy Unit has undertaken a literature review of approaches used by other Government departments, Non Governmental Organisations and private sector to evaluate and monitor policy influencing activities. In addition, ten areas within DFID undertook a pilot to 'projectise' their policy dialogue and influencing activities. Evaluation of the pilot confirmed the value of applying DFID's usual "logical framework" approach to such work, but also noted limitations.

Multilaterals

3.12 In 2006/07 DFID channelled £2,126 million through multilateral organisations¹. That is 43% of their total budget and an increase of 43% from 2002/03 (Figure 8).

Figure 8: DFID multilateral expenditure



Source: Statistics for International Development 2004 to 2007, DFID

¹ Multilateral aid is defined as aid which is channelled through international bodies for use in or on behalf of aid recipient countries. Aid channelled through multilateral development institutions is regarded as bilateral where DFID control the use and destination of funds. This relates mainly to humanitarian assistance delivered through United Nations agencies.

- 3.13 DFID is committed to making funding allocations to multilateral organisations more performance related. It is progressing a joint approach to monitoring and evaluating the performance of multilateral organisations through the Multilateral Organisations Performance Assessment network (MOPAN). Each year MOPAN survey members' country teams in 10 countries about the performance of around three multilateral institutions. The survey centres on issues surrounding partnership working, but also addresses other dimensions of quality. Currently DFID deem this information to be too subjective, and too limited in coverage, to use as a basis for resource allocation, but are in discussion with other donors to make the process more transparent through wider engagement with development partners and introduction of a balanced scorecard approach. In principle, feedback from a wide network of donor bilateral country teams seems a strong source of expert opinion on in-country multilateral performance which is currently under-used.
- 3.14 Information from MOPAN is used to inform Multilateral Development Effectiveness Summaries (MDES), which represents an evolution from DFID's first attempt to assess multilateral performance (the Multilateral Effectiveness Financing Framework, MEFF), drawing inter alia on the findings from a 2005 NAO review of multilaterals. MDES uses a scorecard with the following main assessment areas:
- Building for the future
 - Managing resources
 - Partnerships
 - Country/global performance
- 3.15 The system improves on MEFF by looking for hard indicators of achievement under the various headings, as opposed to focusing on the existence of nominal systems. But it still does not represent an assessment of development impact achieved – largely because the multilaterals themselves have weak systems, and weak reporting, in this area. DFID rate on a three-point scale the multilateral's performance on individual indicators within the areas. They use the results as one input to their multilateral financing decisions. That is one reason why the funding to multilaterals that DFID rates highly, such as the World Bank, has increased significantly.
- 3.16 For an individual multilateral institution, DFID prepares Institutional Strategies (IS) to set out how the UK Government can work with multilateral institutions to achieve poverty reduction, and identify priority objectives and SMART targets to measure progress. Our review of a sample of documents found that IS have not been produced

for all relevant multilaterals, and not all those that have been produced have been updated as required. A DFID M&E audit noted that there is no standard format for institutional strategies and although guidance includes an annex on possible IS format, it does not provide as much detail as is currently available for compiling performance frameworks for Country and Regional Assistance Plans.² The audit also noted an absence of SMART targets, a lack of intermediate milestones, limited baseline data and assigning of responsibility for monitoring the framework or individual indicators. EvD evaluation of DFID multilateral funding programmes, and achievement of relevant IS objectives, is not institutionalised as it is for country programmes.³

- 3.17 DFID has increased its focus on multilateral performance in recent years, and has continued to develop associated assessment tools. Its moves to secure joint assessment with bilateral partners are welcome. Current assessments, however, are still very generalised, both about the performance of a multilateral and about DFID sponsorship of that institution. There is scope to achieve greater precision by continuing the development of current tools, extending the use of customer or DFID country team feedback on multilateral performance (where multilaterals provide insufficient data about their performance), and sharpening the strategies DFID use to guide their funding and monitoring of multilaterals.

² The Baseline Audit of the State of Monitoring and Evaluation in DFID, Roger Drew and Rachel Albone

³ *ibid*

Part Four: Programme Evaluation

- 4.1 Evaluation⁴ can play a significant role in contributing to effective performance management. It can strengthen accountability, promote lesson learning and provide reliable information for management to use in decision making in relation to policy formulation and programme implementation.

Expectations are higher than ever before with regard to the knowledge, evidence and expertise that evaluation should deliver in support of policy decision making and resource allocation.⁵

- 4.2 DFID have a recognised track record and commitment to evaluation as a means of investigating questions on aid effectiveness.

As we spend more on reducing poverty it is vital that we be increasingly rigorous in assessing impact and directing resources to where they can make the most difference to the lives of poor people.⁶

Evaluation Policy and Function

- 4.3 DFID aim for evaluation to be undertaken at every level of the organisation – most projects and programmes undergo some kind of evaluation, for example through an Annual Review or Project Completion Report. DFID do not keep a record of all evaluations conducted across the organisation. Evaluation Department (EvD) undertake a 3 year programme of evaluations (approved by the Independent Advisory Committee on Development Impact), and have the policy lead on evaluation standards and practice, and liaise with other evaluation units. While EvD provides guidance and offers support on request, it does not currently have a strong role in policing evaluation standards across DFID. EVD's role in quality assurance of evaluations across DFID is expected to increase under the forthcoming new Evaluation Policy, which aims to drive up the quality of evaluation throughout DFID.

⁴ DFID define evaluation as the systematic and objective assessment of ongoing or completed aid activities, their design, implementation and results which the aim of determining the relevance, fulfilment of objectives, development efficient, effectiveness, impact and sustainability.

DFID: Guidance on Evaluation and Review: July 2005

⁵ Agency Francaise de Development: A Comparative Study of Evaluation Policies and Practices in Development Agencies: ODI: December 2007

⁶ DFID Annual Report 2008

- 4.4 EvD conducts approximately five Country Programme Evaluations each year with the requirement that all PSA countries are covered over time. In addition, EvD also conducts policy evaluations, for example on thematic activities such as on HIV/AIDS or gender, and major aid instruments such as technical cooperation and budget support. The strategic and thematic reviews, as opposed to specific project evaluation, have raised the profile of EvD within in the organisation. Evaluation is now seen as being ‘central to evidence based policy development.’⁷ The creation of the Independent Advisory Committee on Development Impact is evidence that DFID is committed to an independent and high quality evaluation of international development assistance.

Example of shift in importance of evaluation

‘It is changing. Evaluation is becoming more of a priority and more vigorous by thinking through linkages between what is happening on the ground and strategic objectives.’

‘There has been a shift in the political climate – the need to justify to others and to ourselves what we’re doing.’

Source: NAO: Interviews and focus groups with project teams in Ghana and Sierra Leone

- 4.5 We surveyed DFID staff and other stakeholders about aspects of DFID evaluation - fuller details are at:
http://www.nao.org.uk/publications/0708/nao_survey_perceptions_of_dfid.aspx
79% of staff respondents stated that they were familiar with the role and mandate of EvD. The possible uses for evaluation rated most highly were:
- synthesising lessons from evaluation for use by DFID staff
 - evaluating DFID policies to provide accountability
- 4.6 Accountability and lesson learning functions of evaluation are not mutually exclusive, but they might indicate different priorities for the topics, timings and presentation of evaluations. Such tensions need to be addressed in the organisation’s evaluation policy. There is no finalised policy document detailing the evaluation mandate and function for the organisation, but a policy is under development in collaboration with IACDI, scheduled for formal adoption by March 2009.

⁷ DFID Departmental Report, 2006, page 202

Evaluation independence

- 4.7 There is a tension between evaluation independence and integration with other functions of the organisation.⁸ EvD is not structurally independent, but it is located outside DFID's operational and policy divisions, and like the Internal Audit Department, reports directly to the Director-General, Corporate Performance on the Management Board. 83% of respondents to our DFID staff and stakeholder survey rated the importance of the independence of the evaluation function as being extremely or very important. In terms of the actual independence, only 25% of staff and 23% of other stakeholders rated EvD as entirely or very independent. The key threats to high quality independent evaluation as identified by DFID staff and stakeholders were:
- inadequate resources devoted to evaluation; and
 - organisational pressures to water down critical reports.
- 4.8 Evaluation staff view DFID senior management as supportive of monitoring and evaluation, although such support could be applied more consistently.

Perception of relationship between senior management and evaluation

'It could be a more transparent line [to the MB] to ensure ongoing and direct accountability.'

'We have access [to the MB] but it is not systematically formulated.'

'Yes, they are supportive.....but there is a gap between rhetoric and influence of evaluation work. The message needs to be more consistent especially at the country level.'

'The Management Committee is engaged on an ad hoc basis.'

Source: NAO: Interviews with DFID HO Teams

- 4.9 DFID is not the only organisation seeking to address this challenge. In other aid organisations responsibility for evaluation often lies within a specialised unit managed separately from the rest of the organisation, reporting directly to the Board.

⁸ Agency Francaise de Development: A Comparative Study of Evaluation Policies and Practices in Development Agencies: ODI: December 2007

Agency	Location of Evaluation Department
DFID	Reports to Director General on the Management Board
SIDA: Swedish International Development Co-operation Agency	Secretariat reporting to Board of Directors
DANIDA: Ministry of Foreign Affairs Denmark	Report to Secretary of State outside the operational/management structure
KfW German Development Bank	Report to Board of Managing Directors, outside the management structure
International Monetary Fund	Reports to Executive Board but works at 'arms length'
World Bank	Reports to Board of Directors through Committee on Development Effectiveness (CODE)
African Development Bank	Reports to Board of Managing Directors outside the management structure
European Union	Reports to three Commissioners, administratively attached to EuropeAid

Source: Agency Francaise de Development: A Comparative Study of Evaluation Policies and Practices in Development Agencies: ODI: December 2007

Resources

4.10 EvD contracts-out virtually all its evaluations, using an overall budget of £4.3 million for 2008/9. The budget has increased from £3.6 million in 07/08, and is set to increase further to £5.1 million in 09/10. EvD has 20 staff, who are predominantly involved in tasks such as evaluation programming, commissioning, tendering, quality assurance and management as well as dissemination into DFID and international policy and programming, and evaluation capacity building.⁹ Evaluations which have a joint element, or are of clear benefit to developing countries, can be classified as “programme expenditure”, otherwise evaluations represent “administrative expenditure”. The significance of this distinction is that, with DFID subject to administrative cost reductions as for other Government Departments, the budget for evaluation does not grow automatically with growing programme spend. Non-EvD evaluation work is not centrally recorded, and so total evaluation resourcing in DFID cannot be readily identified.

⁹ Agency Francaise de Development: A Comparative Study of Evaluation Policies and Practices in Development Agencies: ODI: December 2007 p.25

Evaluation Outputs and Quality

- 4.11 EvD have increased their output since our 2002 study, now producing around 24 studies or papers a year. Since evaluations produced outside EvD are not identified or listed, we cannot be sure of the full range of evaluative work that DFID undertakes. But as an indication of the likely scale, our survey respondents identified 30 or so non-EvD evaluations relevant to a two-year period.
- 4.12 One of the issues that affects the value of evaluation is the extent to which findings and recommendations reach staff. Clearly there are a number of indirect routes which can apply – for example, through compliance with revised policy or practice guidance. At the same time, awareness of the source is also important, particularly to get value from evaluation findings, as opposed to broader recommendations. We surveyed a sample of DFID professional and programme staff about evaluations produced in the last two years.
- 4.13 Awareness of individual EvD evaluations varied widely, from 67% to 10% of those responding. There was no obvious pattern in levels of awareness. Looking at the number of evaluations any respondent was aware of yields the following.

Number of Evaluations	% Aware of Existence, Understanding Main Findings or Full Text Amongst DFID Staff	% Aware of Existence, Understanding Main Findings or Full Text Amongst DFID Stakeholders
None	10.6%	43.9%
Between 1 and 5	34.8%	9.8%
Between 6 and 10	21.2.%	17.1%
Between 11 and 20	21.2%	14.6%
Between 21 and 47	7.6%	14.6%
All 48	4.5%	0%

Source: NAO Survey of DFID Staff and Stakeholders

- 4.14 Depth of awareness was relatively low: for every 6 instances of simple awareness of an evaluation, there were only 2 instances of awareness of main findings and 1 of the full text.
- 4.15 We also asked about the quality of presentation. 50% of staff and 67% of stakeholders we surveyed rated the presentation of evaluations as excellent or very good. Our own

views are somewhat less favourable. Interviews with staff during our field visits also confirmed that the “academic” style of presentation put many staff off.

User Friendly Evaluations?

Evaluation of DFID’s Policy and Practice in Support of Gender Equality and Women’s Empowerment: Country Case Studies; India, Nigeria and Western Balkans

Published in 2006, this report was 260 pages long. In addition, the results of the three country offices were not brought together to identify common strengths and weaknesses.

Thematic Studies: Gender Equality through Justice and Rights based policies

The executive summary was very dense making it hard to understand the main findings. In addition, the report was inconsistent in that different sections of the report were written with a very different approach, limiting the ability to compare and contrast different approaches.

General Budget Support joint Evaluation: all volumes together total 1200 pages, with even the synthesis volume running to 250 pages. Very little use of charts, diagrams or graphics to help the reader. However, a series of policy briefs was widely disseminated.

- 4.16 Our survey also asked about the rigour of evaluations: 60% of staff and 78% of stakeholders stated the analysis was fully or generally rigorous. No respondents stated that they found the evaluations unreliable. Our review of a sample of evaluations found that although the methodology adopted in each evaluation appeared to be appropriate and robust, there was limited information presented on methods. The evaluation standards to be used were discussed in only three of the ten evaluations reviewed, of which only two adequately considered, openly presented and explicitly justified their choice of standards. None of the evaluations referred to common evaluation standards other than the DAC measures mentioned by five of the reports.
- 4.17 Our review also found that the degree to which primary evidence was summarised in the report varied widely. Where the evaluation was based on statistical evidence, the recommendations tended to be well substantiated. By contrast, where the evaluations were based on interviews and qualitative research, they often failed to present evidence, quotes or references to substantiate the findings. The recommendations provided were usually appropriate, but often highly generalised in ways not easy to translate to operational circumstances.

- 4.18 In terms of enhancing the quality of evaluation work, the main factor identified by DFID staff and stakeholders was improving the evaluability of projects, programmes and strategies at the outset through clearer and more measurable objectives. Another issue raised was the need to avoid too much emphasis on process, as opposed to results.
- 4.19 Views on joint evaluations – which potentially reduce net burdens or add analytical power, and are in accordance with the Paris Declaration of Aid Effectiveness – were mixed.

Perspectives on joint donor evaluations

‘Harmonising (with other donors) is very time consuming and not easy. But it does help in being influential,’

‘The impact is limited as to ensure agreement (between donors) they feature the lowest common denominator and watered down recommendations. However, the potential impact for lesson learning is huge.’

‘Very time intensive and has enormous transaction costs unless designed (well)’

‘Satisfying multiple donor demands can easily dilute quality of study and comparability’

‘we have to sacrifice rigor of a study to achieve harmony’

‘Help to generate a much better mutual understanding of capabilities, objectives and potential pitfalls in cooperation’

‘I find it very useful as it exposes me to different ways of working and thinking’

‘Very useful in terms of joint learning, influencing and being influenced’

Source: NAO: Interviews with DFID staff at HQ and in country offices

Uses of Evaluation

Potential evaluation benefits as identified by DFID staff	Potential evaluation benefits as identified by DFID stakeholders
More cost-effective aid projects and programmes	Increased institutional learning
Better DFID aid policies	Better institutional accountability
Better aid practices for the international development community	

Source: NAO: Survey of DFID Staff and Stakeholders

- 4.20 52 per cent of staff respondents to our staff survey indicated that they used evaluation material to inform the context for decision making, as opposed to less than 20 per cent using them to inform specific decisions. Evidence from interviews and focus groups conducted during our visits to county offices suggest that staff use informal learning

networks and internet searches to identify good practice and lessons learnt. DFID evaluations may therefore be accessed through these routes. But it was clear that DFID evaluations were not often the first thought staff had when looking for aid effectiveness or good practice information.

Source of good practice and lesson learning information

‘There is a donor gateway on the Internet to access each others information’

‘For lesson learning, I would search for a report from DFID or the World Bank or ODI, or contact another country team.’

‘Information is obtained through team discussion, contacts, networks and by outsourcing research.’

‘It is easier to look on Insight or contact people in other country offices.’

Source: NAO: Interviews and focus groups with project staff in Sierra Leone and Ghana

- 4.21 Our review of a sample of DFID evaluations found them to be more influential at policy level than at field operation level. Where evaluations included country case studies or field work, we found that the analysis and conclusions related to specific countries were not used by field offices as much as one might have expected. On country programme evaluations we were told that country heads and senior management teams valued the chance to get a “peer review” of their programme, but that usually they added few new insights, and so the recommendations did not have great impact. The evaluations therefore had a very limited shelf life. However, progress has been made through annual ‘Holding to Account’ meeting between Directors and Senior Management. The meeting will include challenge over the extent to which findings and recommendations of evaluations have been actioned. And in December 2007, the Evaluation Department introduced an on line tracker which records the recommendations from each evaluation study, the assigned responsibility, the planned response and progress on implementation.

Limited use of evaluation

‘We have attached very little importance to evaluation over the years.’

‘We had a Country Programme Evaluation in 2005 but don’t pay too much attention to it.’

‘The Evaluation Department have limited engagement with country offices, resulting in limited clout to implement recommendations.’

‘The Evaluation Department does not have a high profile, there is confusion over its role and it is regarded as a burden at the field level.’

- 4.22 Evaluation is not explicitly linked to the Public Service Agreements or directed to areas where performance is thought to be particularly weak or strong. There is little evidence to support the integration of evaluation material into reported performance information, or performance management.

“Lessons learnt from evaluation studies do not always feed into decision making at the right time, or in a systematic fashion.”¹⁰

Integration of Evaluation with Performance Management

‘It is difficult to interact ad hoc evaluations with performance management. The balance of the portfolio is not well constructed.’

‘Evaluation should be part of the suite of performance measurement tools’

‘By increasing the audit trail of evaluation work, it could be used as a management support for decision making’

‘There is a gap between understanding the value of monitoring and evaluation and acting upon it. The Management Board need a steer on what to do with monitoring and evaluation information - how to respond to ensure greater accountability and decision making based on results.’

Source: NAO: Interview with DFID staff

- 4.23 There is a clear appetite to undertake quick evaluations to promote lesson learning and ensure findings are useful for the country teams. There is also interest in more impact evaluation undertaken some time after project or programme completion. In recent years EVD has increased its focus on impact evaluation and is playing a leading role

¹⁰ Benchmarking Study on Evaluation Policies and Practices: Full Case Study DFID: ODI p.10

with the World Bank and others, in building international systems and partnerships for impact evaluation. The Head of EVD currently chairs the main international network on development impact evaluation (Network of Networks on Impact Evaluation NONIE).

Example of demand for impact evaluation

‘It is a DFID requirement to undertake a review within six months to identify early impacts. This is not systematic and relies on anecdotal evidence. It would be more useful after 2 years to see what has really changes and learn lessons.’

‘It would be useful to undertake evaluations after 12 months – to see if impacts have been achieved.’

Source: NAO: Interviews and focus groups with project teams in Ghana and Sierra Leone

Part Five: Project and Programme

Monitoring

5.1 Projects and programmes represent the base level of DFID activity, particularly on bilateral expenditure. Performance measurement at this level needs to be adequate to inform management of the project in question, and also feed into higher level performance management frameworks.

Time and resources

5.2 The time spent by staff on monitoring varies according to the type of project and stage of project implementation, and it is not formally distinguished from other duties. DFID country team adviser and programme staff we interviewed estimated that they spent between 10 and 80% of their time on monitoring, with most estimates falling between one quarter and one third. All three teams noted the pressure on monitoring arising from reducing administrative budgets.

Examples of time spent undertaking monitoring and evaluation activities

‘About 20 – 25% but it is a key part of my work.’

‘It is difficult to say as it is a continual process including meetings with the Government and other donors.’

‘Roughly 25% on specific reporting such as Output to Purpose Reviews and Project Completion Report but more time on discussions with partners.’

Source: NAO: Interviews and focus groups with project staff in Sierra Leone and Ghana

Project Indicators

5.3 The findings of our review of project documentation corroborate those of the Baseline Audit of the State of Monitoring and Evaluation in DFID and the Agulhas report *Assessing the Quality of DFID’s Project Reviews*: that there has been a proliferation of indicators (partly in response to the limited availability of data on results), and that many project logical frameworks and associated reviews have not been properly drawn up.

Baseline Audit of the State of Monitoring and Evaluation in DFID Findings on Indicators

- Each of the 115 projects reviewed had between 3 and 64 indicators, averaging 23/project
- 92% of indicators were not disaggregated by or sensitive to gender
- 61% of indicators were process based
- 15% of indicators were based at the output level
- In terms of the input- activity-output-purpose-goal hierarchy, 60% of all project and programme indicators were at lower level than that they purported to represent

Source: Baseline Audit of the State of Monitoring and Evaluation in DFID: Roger Drew and Rachel Albone

5.4 These findings indicate that project monitoring often starts from a flawed base. In particular, the over-representation of process indicators may hinder accurate monitoring of output and goal level attainments. They also signal weaknesses in DFID technical skills and management review processes.

Sources of Performance Information

5.5 DFID staff use a variety of data sources to obtain performance information.

Data Source	Percentage of Log Frames Data Source Quoted in
Document Review	91%
Interview and Discussions	60%
Direct Observations	24%

Source: Baseline Audit of the State of Monitoring and Evaluation in DFID: Roger Drew and Rachel Albone

5.6 The management of performance and evidenced based decision making is dependant upon the availability of high quality and timely data.

"Development talk is full of statistics but they are often old and unreliable and thus fail to capture the policies that produce progress or failure until it is much too late. Badly informed decisions particularly affect poor people."

Rt Hon Clare Short, Secretary of State for International Development, November 1999

5.7 DFID draws its poverty data largely from national governments and multilateral organisations such as the UN and World Bank. To improve the availability and quality

of data DFID primarily supports the strengthening of statistical and data collection capabilities of national Governments, often through international mechanisms, as opposed to undertaking its own data collection.

Examples of DFID support for capacity building

Sierra Leone

DFID support Statistics Sierra Leone in undertaking the collection, analysis and dissemination of accurate and timely statistical information on social, demographic, economic and financial activities to serve the needs of users. Outputs include National Household Survey, Census and Annual Statistical Digest.

Ghana

DFID support the Government of Ghana in developing an informed decision making process by using information from the Ghana Living Survey, Census and Public Expenditure Tracking Survey.

Source: NAO interviews and focus groups with senior country staff, Sierra Leone and Ghana

- 5.8 Relying on locally-run data systems, however, can also bring problems of reliability and of perception.

Examples of data challenges

‘Obtaining data is very difficult, it is not aggregated and contradictory. We don’t know what is happening.’

‘Data can just confirm what we have already learnt through intelligence. If it is contradictory we dismiss the data as there are usually errors in it.’

Source: NAO Interview and Focus Groups with project staff in Ghana and Nigeria

- 5.9 Data capacity building is not incompatible with DFID undertaking its own data collection. During our field visits to DFID country offices, we found examples where DFID had produced, facilitated or verified data collection. For example DFID Sierra Leone had undertaken a service perception survey to obtain the views of the end user and DFID Ghana had commissioned Voices to obtain the view from civil society and build up demand for public services.
- 5.10 Direct observations and field visits are a valuable method to understand the impact of projects and programmes. Interviews and focus groups with country staff revealed that staff undertake field visits from once or twice a year to once or twice a month. But staff also said that given pressures of other work it was difficult to get out as much as they felt would be ideal.

Examples of responses when asked the value of field visits

‘It is an interesting and effective way to understand impact - talking to beneficiaries rather than hearing information through third parties.’

‘Going out can confirm and inform your knowledge of the project. You realise and appreciate much more from the ground than what is reported in the documents. For example: ten beneficiaries are reported on paper, but on the ground you can see what effort has gone into reaching those beneficiaries.’

‘Documents only capture intended benefits. A wider perspective is gained through field visits.’

Source: NAO: Interviews and focus groups with project staff in Sierra Leone and Ghana

Project scoring and performance management

- 5.11 At least annually, DFID staff must assess project progress on a five point scale. The first two points, where the project is wholly or largely on track to meet its objectives, are taken to represent success. The next point – the project is likely to partly meet its objectives – requires the project staff to prepare a remedial action plan. Worse assessments call into question the future of the project.
- 5.12 DFID has recently introduce a new overall measure of portfolio quality which weights the scores attached to each project with reference to a sliding scale (wholly on track = 100%, largely on track – 75% etc) . This change is designed to give fairer weight to performance achieved, and to mitigate the risk of bias in scoring of a sharp cut-off. Nonetheless, quality of project portfolio remains a targeted corporate indicator. And there are clear risks of incentivising staff to score projects optimistically. Staff we spoke to had varying views on the extent to which such risks had materialised.

Views on incentives to score projects highly

‘There is no incentive to score projects highly’

‘There is no pressure. I would rather give a bad mark if that accurately reflected the state of the project.’

‘In cases where I do know the project, I have gone back to the reviewer and asked why have you given this project a 2 when it is a 3? The response is because is it causes extra work.’

‘There is an incentive not fail a project as it causes more work! For example in undertaking six monthly reviews.’

Source: NAO: Interviews and focus groups with project staff in Sierra Leone

- 5.13 More generally, poorly specified log frames create challenges for rigorous project scoring. For example, in the absence of clearly specified targets it is difficult to assess

the extent of progress against expectations. In the projects we reviewed there were up to 33 purpose and output indicators, and many more activity indicators, many of which were qualitative. And while project scoring was reasonable, there was no clear hierarchy or model that helped interpretation of different rates of progress on different indicators. So judgements were subjective, and projects near scoring boundaries could plausibly be scored in either category. DFID has introduced a refinement to the scoring process by asking project to weight individual outputs according to the estimated overall impact on project success. But there is currently insufficient data to determine the effectiveness of this development. DFID's plans to improve the quality of log frames which includes greater specification of baselines, indicators and targets, underpinned by more rigorous economic appraisals, should enable project scoring to be more objective.

Use of Performance Information

Link between performance information and the decision making process

'Performance information is the single most powerful driver, contributing to allocation of staff and decision making process.'

'Too much time and resources has been invested in obtaining and collating performance information for nothing to be done with it.'

Source: NAO: Interviews with DFID Teams

- 5.14 During our visits to DFID country offices, we found several examples of projects that had been revised, documentation redrafted or closed, drawing on monitoring information. Sometimes, the decision flowed from information showing inadequate project progress – as with a Parliamentary support project in Sierra Leone, which was terminated. In other cases, progress had been good, but could not be adequately associated with DFID country objectives or lasting impact - as with projects in Nigeria relating to wetlands and health facilities, which were not continued. These decisions drew on information from a variety of sources, and the extent to which the formal project monitoring data drove the decision varied widely. In the cases we observed, broader staff perceptions of the project in question, the country context and any change in DFID priorities were at least as important as the formal monitoring data.
- 5.15 Staff also wanted readier access to corporate experience in the design and operation of monitoring. While advisers and programme staff had similar views, the latter group in particular felt distanced from broader DFID practice.

Demand for lesson learning

“There is a need for more intensive feedback on lesson learning. It doesn’t happen - we reinvent the wheel every time.”

“Monitoring and evaluation should also focus on lesson learning. For example the success or failure of a flagship project can inform the next programme.”

“If we are going to be working in similar areas, reviews inform future programme and identify lessons in what not to do.”

Source: NAO: Interview and focus groups with DFID staff in Sierra Leone

Training and Guidance on Monitoring and Evaluation

- 5.16 For performance information to be meaningful, DFID needs its staff to structure monitoring log-frames appropriately, and then capture and accurately assess performance data on a timely basis. DFID’s M&E review noted problems with log-frame design, and our interviews and focus groups with country staff showed that many thought they lacked training in M&E – either in general, or, for some mid-career recruits, in specific DFID M&E procedures.

Perception of limited training in monitoring or evaluation

‘Not specifically – I picked it up as I went along’

‘I have not received any training.’

‘I received the two day Project Cycle Management Training but it did not cover monitoring and evaluation in detail.’

Source: NAO: Interviews and focus groups with project staff in Sierra Leone and Ghana

- 5.17 Although 55% of staff surveyed had received Project Cycle Management Training, there is only a small section on monitoring and evaluation and limited guidance on the practical application of log frames and indicators. There is a growing recognition that monitoring and evaluation is an important and valuable part of the work of project staff. 82% of respondents to our survey of DFID staff would value further training on this area. Of those respondents, 72% stated that they would prefer additional training to be organised centrally. DFID plans to roll out log frame training from early 2009.

Views on the value of further training on monitoring or evaluation

‘I would find training useful, especially on the development of log frames to be able to identify issues earlier.’

‘Monitoring and evaluation is a key part of our work, we need to be smarter at it.’

‘Training has had a good impact on my work as I was not sure what I was doing before.’

Source: NAO: Interviews and focus groups with project staff in Sierra Leone

- 5.18 DFID’s M&E review recommends DFID to assess the extent to which staff have monitoring and evaluation competencies required for their work and whether training activities are in place to bridge any gaps.¹¹

¹¹ The Baseline Audit of the State of Monitoring and Evaluation in DFID p.19

Appendix 1

DFID's Public Service Agreement 2003-06

Aim: *Eliminate poverty in poorer countries in particular through achievement by 2015 of the Millennium Development Goals*

Objective I Reduce poverty in sub-Saharan Africa

Target 1 *Progress towards the MDGs in 16 key countries in Africa* **Amber***

Objective II Reduce poverty in Asia

Target 2 *Progress towards the MDGs in 9 key countries in Asia* **Amber ***

Objective III Reduce poverty in Europe, Central Asia, Latin America, the Caribbean, the Middle East and North Africa

Objective IV Increase the impact of key multilateral agencies in reducing poverty and effective response to conflict and humanitarian crises.

Target 3 *Improved effectiveness of international system as demonstrated by:*
- a greater impact of EC external programmes on poverty reduction, including through working for agreement to increase the proportion of EC oda to low income counties from 38% to 70%; and
- ensuring that three-quarters of all eligible HIPC countries committed to poverty reduction receive irrevocable debt relief by 2006 and work with international partners to make progress towards the United Nations 2015 Millennium Development Goals by 2006. (Joint target with FCO)

Amber**

Target 4 *Secure agreement by 2005 to a significant reduction in trade barriers leading to improved trading opportunities for developing countries. (Joint target with DTI and FCO)*

Red ***

Objective V Develop evidence based, innovative approaches to international development

Value for Money

Target 5 *Increase the proportion of DFID's bilateral programme going to low income countries from 78% to 90% and a sustained increase in the index of DFID's bilateral projects evaluated as successful.*

Green***

Source: *Development: Making It Happen DFID 2008 Annual Report, **DFID Annual Report 2007: ***2006 DFID Autumn Performance Report,

Appendix 2

Assessment of the 2003-06 Red and Amber Targets

Target	Assessment	Explanation
One	Amber	Time lags in receiving data mean that DFID are unable to provide a final assessment. Mixed picture for Africa, DFID have reached their target of primary school enrolment, but are unlikely to meet the target for the ratio of girls and boys attending school. This target has been taken forward by Target One of the 2005-08 PSA targets.
Two	Amber	The general picture for Asia is promising with six out of nine sub-targets assessed as on course. This includes a reduction in the number of people living in poverty, an increase in primary school enrolment and a reduction in under 5 mortality rates. Sub-targets which aimed to reduce the prevalence rate of HIV and TB and increase the proportion of births attended by a skilled birth attendant are not on course. This target has been taken forward by Target Two of the 2005-08 PSA targets.
Three	Amber	Most of the sub-targets have been met; the Highly Indebted Poor Countries Initiative (HIPC) was successful; concrete progress was made on EC reforms; and clear steps were taken in support of Poverty Reduction Strategy processes. The target for the proportion of EC funding going to low income countries, however, was not met. This target has been taken forward by Target Three of the 2005-08 PSA targets.
Four	Red	This target was not met due to the failure of the Cancun Ministerial in September 2003, which meant that there was no likelihood of the Doha Development Agenda being completed by 1 January 2005. This target has been taken forward by Target Four of the 2005-08 PSA targets.

Source: DFID Autumn Performance Report 2006, Development Making It Happen 2008 Annual Report

Appendix 3

Methodology

This study used a variety of different methods to gather evidence on key issues:

Country Examinations

We visited Nigeria, Ghana and Sierra Leone in order to understand the extent and use of monitoring and evaluation at an operational level. During our visits we held meetings with staff members and reviewed a sample of project documents.

Assessment of Thematic Activities and Multilateral Funding Instruments

We reviewed DFID's gender and trade thematic activities as well as multilateral funding decisions. This involved holding meetings with key staff members and reviewing relevant literature.

DFID Evaluation Function

We undertook a high level review of DFID Evaluation Department (EvD), looking specifically at how evaluation fits into the performance management framework. We also reviewed a sample of recent evaluations in more detail to establish approach and methods employed and follow up recommendations.

Perceptions Survey

We undertook staff and stakeholder surveys to identify their opinions on the quality and relevance of DFID's evaluation work.

Other Methods

We undertook a literature review of available material on aid monitoring and evaluation, including those undertaken by DFID Internal audit, EvD and performance reporting documentation, as well as the recent monitoring and evaluation review. Our literature review also included an assessment of how other international agencies approach monitoring and evaluation and how their systems are integrated into management arrangements and operational activities. We also undertook semi-structured interviews with key staff based in DFID's headquarters in London and East Kilbride.

Annex A: DFID Response to NAO report on Performance Management

Introduction

1. DFID welcomes this report and particularly the NAO's recognition of the strong progress DFID has made to strengthen monitoring and evaluation over the past few years. The report is timely, coming at a time when DFID has devoted considerable energy to the results agenda – improving measurement and analysis of the results we achieve to maximise the impact of our development assistance.

2. DFID launched its first Results Action Plan in January 2008 – setting out our plan for action. Then in May 2008 the action plan became a central pillar of DFID's Making It Happen change programme which aims to build a Department better able to deal with the challenges of the future.

3. This short response provides an overview of the progress DFID has made in the last year, both during and since the NAO's review. It also sets out briefly our priorities for further action during 2009/10. Our plans already respond to the NAO's recommendations to a large extent – but we will consider the findings and recommendations further over the coming months and consider whether there are areas where we should adjust our plans going forward.

Recent progress

4. During 2008/9 DFID initiated a major programme of work to ensure that allocation and expenditure approval decisions are informed by strong value for money analysis and that DFID has the means to monitor, measure and evaluate results of all its interventions. We also made strong progress on evaluation, increasing resources available, strengthening independence and developing a new policy and strategy.

Data and Indicators

5. DFID recognises that improvements in the quality of data available from partner governments are fundamental to improving the quality of its own performance management information systems.

6. A key element of our Results Action plan is delivery of improved statistics in partner countries. We worked closely with the World Bank and The Netherlands towards the launch of the new 'Statistics for Results facility' (SRF) at the High Level Forum in Accra in September 2008. The facility aims to provide more coordinated support to statistical capacity building through the implementation of National Strategies for the Development of Statistics.

7. We have also introduced a set of standard indicators at both outcome and output levels which will enable us to aggregate the impact of DFID's aid across countries. These cover key service delivery issues such as the number of teachers and health professionals trained and number of classrooms constructed. We are now in the process of introducing a broader set of

suggested indicators for all sectors. This will help staff to identify the best indicators that are available and appropriate for measuring progress.

8. During 2009/10 we plan to finalise and roll out the set of suggested indicators and further embed the standard indicators across our programmes. We will explore how we can derive the most benefits from both the standard indicators and our new management information system (ARIES) including the scope for driving improvements across programmes and more timely reporting on DFID performance. We will retain our focus on implementation of the new Statistics for Results Facility and have agreed to provide £50m over 5 years to build statistical capacity in five pilot countries.

Performance Frameworks

9. DFID progress on performance management was recognised in its most recent Capability Review finalised in March this year. The review found that “performance management across the organisation... have all improved” and that DFID “has a clear corporate performance framework and a strengthened country planning process”.

10. DFID’s corporate performance framework is designed to strike an appropriate balance between managing the delivery of international development outcomes and meeting UK accountability requirements. The framework is designed around a balanced score-card and was further strengthened this year by streamlining reporting on development outcomes with measures of progress against DFID’s own programme of organisational change. The underpinning corporate performance architecture has proved resilient to a particularly challenging development context over the last 12 months.

11. We have strengthened performance frameworks at country level and 15 country business plans completed in 2008/9 include targeted outputs and outcomes. We also improved DFID’s main project monitoring and performance tool, the logical framework (“the log frame”) to ensure it includes clearly specified baselines and targets for the outputs and outcomes each project is expected to achieve. We published new guidance and trained over 100 staff “champions” in the use of the revised format. These champions are now cascading these skills throughout DFID.

12. We also strengthened performance frameworks related to our multilateral funding arrangements. In 2008, Performance Frameworks were introduced into 4 Institutional Strategies (UNDP, UNAIDS, WHO, UNFPA), linking funding to agency performance and new performance based frameworks were put in place for the African and Asian Development Banks. These frameworks use clearer indicators and targets, and have greater clarity about how the objectives and indicators will be tracked over time. Where possible, targets are drawn from agencies’ own plans.

13. We have also taken steps to rigorously assess our own progress towards strengthening the quality of logframes and economic appraisals. This enabled us to focus and prioritise actions to drive improvements. Going forward it will enable stronger scrutiny of compliance with the new approaches.

14. Over the coming year we will make sure that the changes initiated are fully consolidated across the whole organisation to achieve measurable improvements. We will continue to strengthen guidance, build skills and strengthen the internal network of staff keen to learn from others and share good practice. We will improve project review formats to make them more user-friendly and we expect that, under the new Evaluation Policy, the Evaluation Department will report annually on the quality of these project reviews.

15. We plan to introduce performance frameworks into more Institutional Strategies with key multilateral agencies and ensure the logframe approach is implemented consistently across all

DFID multilateral funding. A review of multilateral resource allocation will enhance the link between performance and allocations of the multilateral budget.

Evaluation

16. In 2008/9 a new DFID-wide Evaluation Policy was developed in close collaboration with the Independent Advisory Committee on Development Impact (IACDI), and with full public consultation. The policy will be published in 2009 and will promote:

- a significant increase in number and quality of decentralised evaluations throughout DFID;
- more joint evaluations with partners, in line with Paris and Accra commitments;
- a more focused programme of in-depth evaluations; and
- establishment of evaluation design at the beginning of programmes and policies.

17. The Independent Advisory Committee on Development Impact (IACDI) has now completed its first year of operation and is fully engaged and adding value to DFID's evaluation function. It has advised on policy development and made its first set of recommendations to the Secretary of State. The Committee is now focusing on a full review of the quality of evaluations in DFID.

18. We have made steps to increase staff awareness of evaluation products and to embed the role of evaluation in the organisation. DFID senior management take a strong lead in championing and sustaining a culture of evaluation and learning. Management Board take an Annual Discussion on evaluation and lead Directors have been designated to drive engagement and follow up for all evaluations. The Head of Evaluation sits on a number of Management Board sub-committees. In November 2008, the Head of Evaluation published his 1st Annual Report, highlighting key findings and challenges. Finally, staff have been involved in setting out the topic-list for the next 3 year programme of evaluations.

19. Going forward, the priority will be to finalise and roll out the new Evaluation Policy based on a five-year implementation plan. This will include implementing Evaluation Department's 3-year work-programme of centrally commissioned studies; taking measures to increase the number and quality of decentralised evaluations and reviews carried out across the organisation; and continued engagement internationally to increase and strengthen impact evaluations.

DFID

May 2009