

d·i·e

Deutsches Institut für
Entwicklungspolitik



German Development
Institute

Multilateral Organisations Performance
Assessment: Opportunities and Limitations
for Harmonisation among Development
Agencies

Andreas Obser

Bonn 2007

Discussion Paper / Deutsches Institut für Entwicklungspolitik
ISSN 1860-0441

Obser, Andreas: Multilateral organisations performance assessment: opportunities and limitations for harmonisation among development agencies / Andreas Obser. – Bonn: Dt. Inst. für Entwicklungspolitik, 2007. – (Discussion Paper / Deutsches Institut für Entwicklungspolitik ; 19/2007)
ISBN 978-3-88985-357-8

© Deutsches Institut für Entwicklungspolitik gGmbH
Tulpenfeld 6, 53113 Bonn
☎ +49 (0)228 94927-0
☎ +49 (0)228 94927-130
E-Mail: die@die-gdi.de
<http://www.die-gdi.de>

Contents

Abbreviations

Abstract

1	Introduction	1
3	Sisyphus' job: measuring performance of multilateral organisations	4
3.1	Multiple purposes: a taxonomy	4
3.2	Diverse practices: an international comparison	7
4	Achille's heel: managing performance in development agencies	14
4.1	Four types of performance management approaches	14
4.2	Donor readiness assessment for multilateral performance assessment	17
5	Some conclusions and remaining challenges	22
	Bibliography	25
	Annexes	31

Boxes

Box 1:	UNICEF's partnership performance at a glance: summary of findings	36
Box 2:	Examples of MfDR in development agencies	37
Box 3:	Example of MEFF: UNIDO	43

Figures

Figure 1:	Relationship between measuring, integrating, and using performance in the evolving models	16
-----------	---	----

Tables

Table 1:	Taxonomy of multilateral performance assessments	5
Table 2:	3 avenues of multilateral performance assessment	7
Table 3:	Four models of performance	15
Table 4:	RBM ideal types and OECD/DAC donor systems	17
Table 5:	Components and pillars of COMPAS	39
Table 6:	The MEFF checklist and major findings	41
Table 7:	Interim MO relevance and effectiveness assessment (MERA) of CIDA	44
Table 8:	Overview of PMF for multilateral development cooperation	49
Table 9:	The "New Approach"	52

Abbreviations

ADB	African Development Bank
BO	Bilateral Organization
COMPAS	Common Performance Assessment System
CIDA	Canadian International Development Agency
DAC	Development Assistance Committee
DANIDA	The Danish International Development Agency
DFID	UK Department for International Development
FAO	Food and Agricultural Organization of the United Nations
IFI	International Financial Institution
M & E	Monitoring and Evaluation
RbM	Results-based Management
MDBs	Multilateral Development Banks
MDGs	Millennium Development Goals
MEFF	Multilateral Effectiveness Framework
MERA	Multilateral Evaluation Relevance and Assessment system
MfDR	Managing for Development Results
MMS	Multilateral Monitoring System
MO	Multilateral Organization
MOPAN	Multilateral Organizations Performance Assessment Network
MPA	Multilateral Performance Assessment (frameworks)
NPM	New Public Management
ODA	Official Development Assistance
OECD	Organisation for Economic Cooperation and Development
PCM	Project Cycle Management
PFM	Public Financial Management
PMF	Performance Management Framework
PSIA	Poverty and Social Impact Analysis
RBM	Results Based Management
SIDA	Swedish International Development Agency
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNDP	United Nations Development Programme
UNFPA	United Nation Population Fund
UNICEF	United Nations Children's Fund
UNIDO	United Nations Industrial Development Organization
WHO	World Health Organization

Abstract

Donors' concern with assessing multilateral organisations' performance reflects an increased focus on effectiveness and development results. The international community is addressing these issues as essential part of the common effort to achieve the internationally agreed Millennium Development Goals (MDGs). The paper proposes a taxonomy of four main purposes why bilateral agencies may consider multilateral performance assessments. Unfortunately no single approach currently practiced is appropriate for all four purposes. It is suggested that bilateral agencies should not seek the one magic all-inclusive approach. While recognising the importance of donors' political and administrative traditions, three types of donor systems are identified: (a) donors that "administrate multilateral performance", (b) donors showing "management of multilateral performances" and (c) donors exercising "multilateral performance management". In practice and with regard to harmonization, a modularised strategy seems to be a functional way to make multilateral performance assessment work for all.

1 Introduction

The ‘results’ agenda in development

In international development cooperation there is an increasing number of international initiatives and forces at work pushing and prodding development agencies in the direction of adopting management systems geared toward reform and above all, results. Since the Monterrey Conference of 2002, the development community – partner countries and donors alike – has been focusing on managing its work so as to achieve the maximum development results. This is not to say that development efforts never achieved results before 2002, but what is new is a shared understanding of the need to think about results, and ways to achieve them, from the beginning of any initiative, enterprise, or project – and then to monitor progress and continue shaping the effort so that the desired results are actually achieved” (cf. OECD / World Bank 2006).

The OECD/DAC Joint Venture on Managing for Development Results (JV MfDR)¹ brings together bilateral development agencies, multilateral organisations and partner countries in order to promote the concept and practice of results-based management to its members and partners and develop shared values, methodological approaches and co-operation procedures. Managing for development results is one of the five basic principles of the Paris Declaration on Aid Effectiveness, signed in March 2005 (MfDR 2006).

Why measure results of multilateral organisations?

Multilateral organisations handle up to one third of global aid volumes and, in some cases, play a unique role in making progress towards the Millennium Development Goals (MDGs). There is a strong case for a growing multilateral system as an effective way to deliver the increased aid flows needed to reach the MDGs. The global partnership for results prompted by the Monterrey Consensus has also been changing the way development agencies work (CIDA 2003). The growing complexity and ongoing advances of the multilateral system make it necessary for bilateral development agencies to adapt and further develop their policy accordingly, and deploy their resources so as to ensure effective and efficient collaboration in multilateral development efforts (SDC / SECO 2005). Many bilateral development agencies are exposed to external and internal pressures for improved management of their engagement with multilateral organisations. Evidence on multilateral effectiveness is being called in from all directions², e.g.:

- to better inform multilateral policy, institutional strategy and financial allocation decisions;
- to better exercise accountability and reporting to national constituencies;
- to strengthen participation in the governance of multilateral organisations (e.g. at Board or Governing Body meetings);
- to strengthen relationships with multilateral organisations at country level;

1 The JV MfDR is a subsidiary body of the OECD/DAC Working Party on Aid Effectiveness. For more information see: <http://www.mfdr.org/>.

2 See CIDA (2006a); DFID (2005b); MOPAN (2006).

- as input to joint advocacy by like-minded bilateral groups such as Utstein;
- as input into wider debates about aid effectiveness.

A major problem has been, that there have been no systematic, comparative, across-the-board assessments of multilateral organisations carried out before, when the issue started to become prominent about 4–5 years ago. More fundamentally, there has been little international consensus on *how* such an assessment should be carried out. There has been little agreement on how to define multilateral effectiveness, far less on how to measure it (cf. DFID 2005a, 6).

Is there influence from the private sector or relevant experience in government in general?

In the history of international aid administration there have been different movements to introduce management instruments that help development agencies to check how well they do by emphasizing targets rather than processes and inputs (i.e. fiscal resources and personnel). The focus is no longer on activities and functions of the agencies but on the results they achieve. Since the 1990s, RBM is part of an effort to introduce private-sector management techniques into development agencies, driven by what came to be known as New Public Management (NPM) (Bouckaert / van Dooren 2003; Bouckaert et al. 2006, 1; CIDA 2003; Pollitt / Bouckaert 2004).

Talbot (2005) emphasises the wider influence from the private sector and in particular the “quality” and “benchmarking” movements. Two well-known generic performance models are the Baldrich Awards (George 1992) and the Awards of the European Foundation of Quality Management (EFQM). Both Baldrige and EFQM models are annual snapshots of performance based on either examination by external assessors or self-assessment. They are meant to form the basis of subsequent improvement programs. One interesting difference between the two is the inclusion criteria for “corporate social responsibility” in EFQM in the form of an “impact on society” cluster. Talbot suggests, this may be due to the differences between the more free-market, liberal democratic traditions of the USA and more social democratic traditions in Europe (cf. *ibid.*, 506).

Over the past two decades there has been a trend within governments across the globe to reform and reshape the ways in which they function. There are increasing demands for government accountability for results, transparency, and provision of more efficient and effective services. Results-Based Management (RBM) has become a key feature in public agencies. By providing information on the degree to which agencies reach targeted outputs and outcomes, RBM introduces a means to check on performance and hold public agencies accountable. The scope, sophistication and public visibility of performance measurement activities are unprecedented (Bouckaert / Halligan 2006; Heinrich 2003, 25).

What is new compared to former management approaches in development?

It is not a new phenomenon that development agencies monitor and evaluate their own performance. Traditional RBM focuses on the monitoring and evaluation of inputs, activities, and outputs; that is, project program implementation. Development agencies have over time tracked their expenditures and revenues, staffing levels and resources, program and project activities, numbers of participants, goods and services produced, and so on. Performance-informed RBM combines the traditional approach of monitoring implemen-

tation with the assessment of results. It is this linking of implementation progress with progress in achieving the desired objectives or goals (results) of development policies and programs that makes performance-informed RBM most useful as a tool for public management (cf. Kusek / Rist 2005, 301).

RBM is different to traditional Project Cycle Management (PCM). PCM – by applying the Logical Framework Approach – is the process that begins with the initial conception of a development intervention and concludes with post-completion evaluation. The PCM still is a common practice among many DAC member countries (OECD 2005a, 86). PCM focuses on inputs and then analysing their potential impact. RBM in comparison focuses on desired development outcomes and identifies what inputs and actions are required to achieve them. It shifts the emphasis to focus on the desired results and encourages flexibility and responsiveness in programming to achieve them.³

What is the role of policy evaluation vis-à-vis performance management and RBM?

The role of evaluation vis-à-vis performance management has not always been clear-cut:

- policy evaluation tended to focus on programmes, to carry out episodic “one – off” studies and use more qualitative methods alongside quantitative ones (although there have been huge disputes about method) and its outputs were aimed more at policy;
- performance measurement had a more organisational focus, tended towards setting up long-term continuous monitoring systems, relied slightly more on quantitative approaches and its outputs were aimed more at management (cf. Talbot 2006).

In part, the differences are because evaluation was well established in many governments before the introduction of performance management and the new approaches did not necessarily incorporate evaluation. A report by OECD (2001a) reveals that new performance management techniques were developed partly in response to perceived failures of evaluation; for example, the perception that uses of evaluation findings were limited relative to their costs. Moreover, evaluation was often viewed as a specialized function carried out by external experts or independent units, whereas performance management, which involves reforming core management processes, was essentially the responsibility of managers within the organization. Failure to clarify the relationship of evaluation to performance management can lead to duplication of efforts, confusion, and tensions among organizational units and professional groups. Most OECD governments see evaluation as part of the overall performance management framework, but the degree of integration and independence varies. Several approaches are possible:

- At one extreme, evaluation may be viewed as a completely separate and independent function with clear roles vis-à-vis performance management.
- At the other extreme, evaluation is seen not as a separate or independent function but as completely integrated into individual performance management instruments.

3 See CIDA (2003) and PARC <http://www.iod.uk.com/>.

- A middle approach – reflected in respective OECD best practice guidelines – views evaluation as a separate or specialized function, but integrated into performance management (OECD 2001a, 8).

Challenges and dilemmas for Results-Based Management (RBM)

For all its benefits, RBM is subject to a number of inherent unwanted dynamics that are difficult to counteract: (i) “creaming”, (ii) “teaching to the test” and the tendency to (iii) set low goals (Bouckaert et al. 2006, 8 ff.)⁴.

Agencies and organizations subject to performance reporting may engage in “creaming” by selecting clients or cases that are easy to treat, thus ensuring high performance results. For example, an organization delivering health services which is evaluated based on the health of its clients may avoid taking on very ill clients. An organization evaluated based on the number of drug addicts treated successfully may avoid treating long-term addicts with poor prognoses for recovery. The same goes for schools. If schools can select their pupils, as private schools do, they have incentives to choose those who will likely perform well.

RBM can lead organizations to “teach to the test” by focusing on those activities that contribute directly to the achievement of performance goals, possibly leaving aside activities that are valuable for the common good but omitted from performance measurement. This dynamic can be observed in schools evaluated based on standardized tests. In order to achieve high performance results, individual teachers or the school as a whole have an incentive to focus on teaching the material that will be tested, while leaving aside possibly valuable skills that are not part of the test. In the health care sector, professionals may have greater incentives to treat symptoms rather than to alleviate the underlying causes of an illness or engage in prevention.

A third unwanted dynamic is the potential tendency of agencies and organizations subject to RBM to set low goals in order to insure high performance. If possible, the agency or organization may avoid setting goals at all. In the field of education, this means that schools underestimate the ability to teach a certain number of students with a set budget (output) or when they underrate the rates of students expected to pass certain tests (outcome). In the health sector, the tendency to set low goals may become visible when health organizations underestimate their capacity to treat a certain number of patients given their resources (outputs) or when they express limited expectations to treat health conditions (outcome).⁵

2 Sisyphus’ job: measuring performance of multilateral organisations

2.1 Multiple purposes: a taxonomy

Multilateral performance assessment is not an end in itself. Why should bilateral development agencies measure performance of multilateral organisations? There are two fun-

4 On trade-offs, balances, limits, dilemmas, contradictions and paradoxes of public management reforms in general see Pollit / Bouckaert (2004).

5 See also Bird et al. (2005).

damental rationales: (i) using the performance information to continuously improve a development agency's own organisation and management – i.e. the *internal purpose*, and (ii) a general need to document and demonstrate a development agency's multilateral performance in and with multilateral organizations to its domestic constituency such as parliament and the general public – i.e. the *external purpose*.

Anyhow, with the up-scaling of the results agenda (e.g., Paris Declaration and the MDGs) the external purposes of multilateral performance assessment become extended beyond national spheres. Partner countries and the international community in general become additional target groups of development agency's multilateral performance measures, leading to at least four key purposes in all (see table 1).

Purposes of multilateral performance assessments	Target group
Multilateral performance assessment as <i>Public Management</i> (a) as <i>output efficiency</i> (b) as <i>allocation efficiency</i> (c) as <i>effectiveness</i>	Development Agency
Multilateral performance assessment as <i>Public Accountability</i>	Domestic constituency
Multilateral performance assessment as <i>Client Service</i>	Partner countries
Multilateral performance assessment as <i>Global Public Value</i>	International community

The four purposes listed here are not all-inclusive and nor is it suggested that all of them appear in every bilateral development agency, simultaneously and equally. On the contrary, they have been erratic, patchy, and far from universal in their emergence with very few development agencies making an effort to adopt them all, several are taking up a selected set and some hardly any of them at all (Behn 2003; Nicholson-Crotty et al. 2006; Talbot 2005, 496 ff.).

Multilateral performance assessment as *Public Management* involves three essential dimensions. These again may be appreciated and connected by each development agency in its own way:

- (a) Improving the *output efficiency* of multilateral cooperation is perhaps the most traditional argument for the use of respective performance measures. It is about administrative control mechanisms inside development agencies (Hood et al. 1999) and about turning multilateral inputs into outputs.
- (b) Striving for *allocation efficiency* is a long-running argument in public management too. It is concerned with rational planning in development agencies' budgeting process. Multilateral performance information, it is argued, is necessary to understanding the utility of resources allocated to multilateral organisations. UK's Public Service Agreements (PSA) are frequently referred to as highly developed applications in this tradition (Chancellor of the Exchequer 2004)⁶.

⁶ See chapter 11, on DFID (http://www.hm-treasury.gov.uk/media/F5D/2D/sr04_psa_ch11.pdf).

- (c) In a way, increased emphasis on *effectiveness* has emerged as a reaction to the efficiency position. The argument shifts the focus to outcomes and impact of development agencies' multilateral policy. It is suggested that development agencies have concentrated on the task of implementing (isolated) multilateral activities efficiently, at the cost of losing sight of the overall aims of development policy – i.e., the eventual outcomes or development results achieved as well as better coordination within the whole organizational system of development agency⁷ (OECD / World Bank 2006; Talbot 2005).

Multilateral performance assessment as *Public Accountability* suggests that general public need to be given information not just about what is spent on multilateral cooperation but also what results are achieved. Development agencies have mixed policies on how much and in what detail information is published. It became common practice however, that development agencies publish annual reports including regular information on multilateral cooperation.

Multilateral performance assessment as *Client Service* is linked to the Paris Declaration which lays down a practical, action-orientated roadmap to improve the quality of aid and its impact on development. The 56 partnership commitments are organised around the five key principles: ownership, alignment, harmonisation, managing for results, and mutual accountability. Addressing these principles and monitoring the related set of 12 indicators of aid effectiveness is linked to notions about better partner services by overcoming lack of coordination and strengthening mutual accountability. This purpose focuses on “excellence” and baselines to be employed in multilateral performance assessments (OECD 2006a; World Bank 2006a).

Multilateral performance assessment as *Global Public Value* is the most recent argument linked to the MDGs. It suggests that multilateral organisations have a more positive role in creating global public value which cannot be made by bilateral development agencies. Considering this added value in multilateral performance assessments usually reflects in general a more positive attitude to multilateral organisations achievements and mandate. This argument is pressed ahead and beyond multilateral organisations as only provides of ‘regional and public goods’ (Kaul 2004; Ferroni / Mody 2002) by addressing the role and performance assessment of global public policies and programs (OECD / World Bank 2007; World Bank 2004).

Unfortunately, no single performance measure is appropriate for all four purposes. Consequently, development agencies should not seek the one magic performance assessment approach. Instead, they need to think about the strategic purposes to which multilateral performance assessment might contribute and how they might deploy these measures. Only then can they select and harmonise measures with the characteristics necessary to help achieve each purpose. Without at least a tentative policy about how multilateral performance measures can be employed to foster effectiveness of multilateral cooperation, development agencies will be unable to decide what should be assessed and how to harmonise procedures (cf. Behn 2003; compare also, Nicholson-Crotty et al. 2006).

7 E.g. IFI- vs. UN- vs. Global Programs-units, bilateral vs. multilateral units, headquarter vs. field offices.

2.2 Diverse practices: an international comparison

The following section and related annexes are for the most part based on findings of 2 recent studies. ODI’s synthesis report on *Relationships of other donor organisations with multilaterals* (ODI 2005) and the survey on *Bilateral Methodologies for Assessing Multilateral Performance*, commissioned by CIDA (2006b). The Canadian survey – very similar to ODI’s study – had two key objectives:

- *Identify donor countries’ current methodologies.* Asking and comparing donors on the importance attached to methodologies, the basic nature of methodology, current flaws in the flow of information originating from multilateral organisations, and the weight given by bilateral donors to various considerations in allocation of resources.
- *Elicit donor countries’ views concerning the future.* Both studies address issues such as the need for consensus among donors regarding indicators, the feasibility of a standard framework in this respect, the type of groupings of multilateral organisations that may be required to facilitate standardization, and similar concerns.

At present, three avenues including eight different approaches of multilateral performance assessment may be differentiated: (i) the joined-up avenue showing approaches involving several donors or development agencies with a focus on harmonisation, (ii) the bilateral avenue including specific approaches adopted by individual donors, and (iii) the complementary avenue having approaches that address perceived deficiencies of both joined-up and bilateral approaches (see, table 2):

- *The Multilateral Organisations Performance Assessment Network (MOPAN)*
- *The Common Performance Assessment System (COMPAS)*
- *The British Multilateral Effectiveness Framework (MEFF)*
- *The Canadian Multilateral Evaluation Relevance and Assessment system (MERA)*
- *The Danish Multilateral Monitoring Survey System (MMS)*
- *The Dutch Performance Management Framework (PMF)*
- *The New Approach*
- *The Client Surveys*

Table 2: 3 avenues of multilateral performance assessment		
Joined-up avenue	Bilateral avenue	Complementary avenue
<ul style="list-style-type: none"> ▪ MOPAN ▪ COMPAS 	<ul style="list-style-type: none"> ▪ MEFF ▪ MERA ▪ PMF ▪ MMS 	<ul style="list-style-type: none"> ▪ New Approach (Client Surveys)

*The Multilateral Organisations Performance Assessment Network (MOPAN)*⁸

MOPAN is a network of nine “like minded” donor countries, which includes Austria, Canada, Denmark, Finland, the Netherlands, Norway, Sweden, Switzerland and the United Kingdom, who share, among other things, a commitment to strengthening and supporting the multilateral system. France is about to join as 10th member.⁹ MOPAN conducts a joint Annual Survey of multilateral organizations’ partnership behaviour with national stakeholders and other donor organizations at country level. Each year, the Annual Survey covers 3–4 multilateral organisations and is conducted in 8–10 countries. A Country Report is produced in each of these countries and a Synthesis Report brings together all the results. The Secretariat function rotates among the MOPAN members annually.

Since 2003 MOPAN has conducted an Annual Survey on selected multilateral organisations in a number of countries where MOPAN members have Embassies or country offices. As a rolling exercise, the Annual Survey should, over time, be able to cover most of the major development. The first Annual Survey was implemented in 2003 as a pilot survey followed by three full fledged Annual Surveys in 2004, 2005 and 2006¹⁰. The Synthesis Reports for these Annual Surveys are available from MOPAN members. The stated objectives are:

- Better information and understanding of multilateral organisations, their roles and performance by decision makers concerned, parliamentarians and the general public in MOPAN member countries,
- better informed dialogue with the multilateral organisations, both at headquarters and at country level,
- engagement of MOPAN country offices in the assessment of multilateral performance and,
- improved overall performance of multilateral organisations at country level.

MOPAN’s assessment is an opinion survey, eliciting the informed judgment of MOPAN member embassy or country office staff. Country-level staff members surveyed are asked to give their perceptions on the performance of different organizations relative to their respective mandates in the countries in which they serve. MOPAN surveys focus on the quality of national partnership (contribution to policy dialogue, advocacy etc); and inter-agency partnership (information sharing, interagency coordination). The final report includes findings from individual questionnaires, agency templates, country team discussion, country reports and a synthesis report.

Members refer to no direct effects on budget decisions as the MOPAN is expected to be used alongside other evidence-based exercises such as internal evaluations by the organisations themselves, and external (e.g. multi-donor) evaluations. Financing decisions are complex and rely on a number of factors.

8 See annexes 2 and 3.

9 Germany, although member of the Utstein Group, not actively participates in MOPAN.

10 The MOs selected for the 2006 MOPAN survey are UNICEF, ILO and ADB. The exercise took place in the following countries: Burkina Faso, Kenya, Mozambique, Uganda, Indonesia, Nepal, Pakistan, Sri Lanka, Colombia, and Guatemala.

Perceptions are useful if they are based on informed judgements, which in this case are formed on the basis of day-to-day interactions with the multilateral organisations. Many MOPAN members work with these organisations through co-financing or participation in joint initiatives and donor meetings. They have an opportunity to observe multilateral organisations' partnership behaviour directly. The survey asks only about behaviour that is observed by MOPAN members. The MOPAN members' joint group discussion provides a mechanism for testing individual views and pooling information. A weakness is referred to some cases in which MOPAN participants may have limited knowledge of the multilateral organisations.

Among the positive spin-offs for members the most important ones are referred to the increased knowledge at country level of the way multilateral organisations operate, the sharing (at country and headquarter level) of experiences in working with multilateral organisations, and the shared understanding at MOPAN headquarter level of each others' policies and practices regarding the multilateral agenda. In an effort to promote transparency and dialogue among stakeholders, MOPAN has asked multilateral organisations to provide their feedback on survey conclusions, and posts it on their websites together with the MOPAN Survey Report.

The MOPAN monitoring exercise is supposed to remain a light and rapid exercise. It is organised to keep transaction costs for all concerned as low as possible, without undermining the validity of the assessments (MOPAN 2006, 39). MOPAN is considered as "the only game in town for effective donor harmonisation". At the same time, the flexible network arrangement (e.g. rotating secretariat roles among members) is seen as potential challenge of "who speaks for MOPAN" when enhancing harmonisation with other initiatives.

The Common Performance Assessment System (COMPAS)¹¹

In the past, international financial institutions have, largely focused on inputs and processes rather than on development outcomes. Moving to a results management agenda will require a shift in institutional practices – which has only just begun with the new efforts to develop a common performance measurement system (COMPAS) and integrating Management for Development Results into multilateral development banks' practices (World Bank 2006c, xix). COMPAS is devised by the MDBs and aims to provide managers and shareholders of the MDBs with a common source of information on how MDBs are contributing to development results.

This first COMPAS report – published by ADB in 2006¹² – provides the initial basis for a framework for collecting consistent and comparable information on how the MDBs are implementing managing for development results (MfDR). The MDBs have succeeded in drafting a set of common indicators that may form the basis of future reporting by them (CIDA 2006b). It is based on seven categories of data: (a) country-level capacity development, (b) performance-based concessional financing, (c) country strategies; (d) projects

11 See annexes 4 and 5.

12 Multilateral Development Banks' Working Group on Managing for Development Results (2006).

and programs, (e) monitoring and evaluation, (f) learning and incentives, and (g) inter-agency harmonization.

Analysis of data in the first COMPAS report suggests overall progress in the implementation of the MfDR agenda. Awareness of results is increasing, and frameworks, systems, and procedures are being implemented in all the institutions. The degree of institutionalization of the MfDR agenda, however, varies across its dimensions and among MDBs. Although the design, approval, and implementation of new procedures and systems at the institutional level is likely to take time, a key challenge for all MDBs will be to go beyond the introduction of such systems and procedures and to establish an institutional culture of using the information on results to inform decision making. Below is a summary of the progress and challenges in each of the three pillars of the COMPAS (cf. World Bank 2006c, 107).

The British Multilateral Effectiveness Framework (MEFF)¹³

During 2003–2004, DFID set up the MEFF for assessing and monitoring the institutional effectiveness of international organizations it funds. The system was developed internally and DFID staff conducted the assessments. Twenty-three organisations have been assessed. MEFF's key objectives include to:

- Provide information to strengthen DFID's accountability under its Public Service Agreement (PSA) objectives
- Provide inputs to DFID's institutional strategies for engagement with multilaterals,
- Provide inputs to future financing decisions

The MEFF focuses on eight corporate management systems (i.e. corporate governance, corporate strategy, resource management, operational management, quality assurance, staff management, monitoring, evaluation and lessons learning, and reporting of results. The measurement system covers three elements: internal performance, focus on country-level results, and partnership. It uses three main assessment instruments, comprising a checklist of indicators, expressed as questions; a scorecard rating data in the checklists; and a summary report, providing a brief, qualitative review of an organization's accomplishments. The framework is designed to be applicable to all multilateral organizations, but recognizes the need to differentiate among different groupings: multilateral development banks, UN development organizations, UN standard-setting organizations, humanitarian organizations, coordinating organizations and the European Commission.

The Multilateral Effectiveness Framework is coming closest to a comprehensive measurement approach. Some respondents of the survey by CIDA (2006b) expressed concerns that the selected groupings of multilateral organisations contain some overlap (e.g. UNICEF is both a development and a humanitarian organization), or may not allow sufficiently for the measurement of individual characteristics.

CIDA's survey study revealed general respect for MEFF among donor countries, and most respondents considered it to be a good starting point for progress towards the building of

13 See annexes 6 and 7.

consensus and a harmonized approach towards measurement. At the same time, MEFF is suspected that it could lead to a somewhat mechanistic approach to funding allocations (ODI 2005).

The Canadian Multilateral Evaluation Relevance and Assessment (MERA)¹⁴

The Canadian International Development Agency (CIDA) has developed a standard template of questions called the Multilateral Evaluation Relevance and Assessment system, MERA. The purpose is:

- To better inform policy and financial allocation decisions on more solid evidence. Ultimately, CIDA aims at getting more predictable multi-year funding, and an adequate balance between core and responsive/earmarked funding.
- To better exercise CIDA's accountability and improve its reporting to Canadian citizens.
- For more effective Board meetings and better identification of areas requiring improvements.
- To deliver on MPB commitment to complete the Multilateral effectiveness review, as part of the Agency Aid Effectiveness agenda (CIDA 2006c).

MERA covers three different themes: relevance, effectiveness and improvement measures in the management. Multilateral organisations will be rated on a scale of 1 to 5 for each of the 3 themes. MERA provides a framework for comparing multilateral organisations receiving core funding within each Directorate. Assessments will be based on the submissions made by Program Managers of the MERAs, supplemented by additional information, such as the institutions reports, the field surveys, any reports indicating results achieved. The next step is said to review the current level of funding to these multilateral organisations and propose a new allocation, as required.

Under "relevance", the measurement includes the links to the MDGs and role in the international architecture. Measurement of improvements relate to those driven by donors or recipients, as well as general improvements. The template is applicable to all multilateral organisations but is considered to suffer from a serious lack of detail. Finally, MERA is regarded as an interim system requiring elaboration and specificity while awaiting a more harmonized and internationally agreed methodology (CIDA 2006b).

The Danish Performance Management Framework (PMF)¹⁵

In 2003, DANIDA developed a Performance Management Framework (PMF). The objectives of PMF are to: enhance the quality of Danish development cooperation through stronger focus on results (Danida 2005a); improve management and continuous learning, through better information and reporting; and strengthen accountability through performance assessments and measurement in the context of an increasingly decentralized management structure. PMF highlights bilateral and multilateral agencies' respective monitoring and evaluation objectives, key management documents, methodology and source of

14 See annex 8.

15 See annex 9.

information and reporting responsibilities at the corporate, country (Danida 1991) and sector levels. PMF enables bilateral and multilateral agencies to track their progress in meeting agreed-upon monitoring and evaluation standards.

The Danish Performance Measurement Framework (PMF) includes a number of tools such as analysis of organizational strategies, perception analyses (own embassies and through MOPAN), and assessments of multilateral organisations' evaluation and reporting systems. In the measurement of both organizational practices and field level results, the PMF relies on the objectives, targets and indicators established by the relevant organizations themselves as set out in the organization's own vision and strategies. (CIDA 2006b; Danida 2005b).

In addition, Danida's Evaluation Department has been piloting a "New Approach" (see below). This new approach reviews the extent to which RBM principles form the basis of the work of monitoring and evaluation units within the multilateral organisations. The approach adopts the principle that measurement of multilateral organisations should be based increasingly upon the evidence and documentation produced by the organisations' own monitoring and evaluation systems (CIDA 2006b).

The Dutch Multilateral Monitoring System (MMS)¹⁶

The Multilateral Monitoring Survey (MMS) is an annual updated performance assessment for the most important UN institutions and International Financial Institutions at country level of the 36 Dutch partner countries. Primary sources of information for the Multilateral Monitoring Survey (MMS) were the Beoordelingskader (BOK) questionnaires of 2002 and 2003 answered by embassy, permanent representation of mission, and constituency offices staff. Staff and peers are asked to comment on a number of specific questions encapsulating corresponding indicators (e.g. good governance, poverty reduction, accountability and financial management, relevance to Dutch policies and strategies, etc.)(CIDA 2006b).

The New Approach¹⁷

In January 2004, DANIDA tabled a proposal to the DAC Evaluation Network for a "New Approach" to evaluating multilateral assistance.¹⁸ The "New Approach" calls for strengthening multilaterals' pre-existing RBM and evaluation systems, with the goal of enabling bilateral donors to depend on multilaterals' own results reporting for performance assessments.

In order to determine whether a multilateral agency's own central evaluation office has produced, or is capable of producing, evaluations that are credible, valid and useable (i.e. comparable to evaluations that are commissioned through an independent multi-donor evaluation), the "New Approach" proposes to first assess: i) the *independence* of a multilateral organization's evaluation systems; ii) *credibility* of the evaluation process; and iii)

16 See annex 10.

17 See annex 11.

18 See Danida (2006a), furthermore a comparable exercise with UNICEF (DFID 2006).

use of evaluation evidence by these agencies, beneficiary countries' and bilateral partners. This assessment is carried out by a Peer Panel. Since the proposal of the "New Approach", DANIDA has conducted one evaluation of the UNDP in 2005 (Danida 2006b; UNDP 2006). CIDA has taken the lead in conducting an evaluation of UNICEF using the "New Approach" in 2006 (UNICEF 2006a; 2006b).

Intermediate summary: comparative strengths and weaknesses of current practices

The CIDA survey sums up that there was virtual unanimity about the key justification for multilateral performance measurement: multilateral organisations' accountability. Aside from accountability, respondents all agreed that performance measurement allowed the identification of strengths and weaknesses as a necessary basis for corrective action. Many mentioned also that measurement is important to maintain public support for the multilateral system. Donors also feel that, although an important consideration in budgetary planning, measurement is not the only criterion in determining the level and type of financial and technical support.

Method of measurement among development agencies vary considerably in sophistication. At least half of respondents reported that they did not have a defined methodology as such, and relied on a "considered weighing of strengths and weaknesses". Two respondents reported the use of specific indicators: (i) the UK's Multilateral Effectiveness Framework (MEFF) which classifies the multilaterals into separate groupings according to functionality and provides process indicators designed for each; and (ii) the Canadian Multilateral Evaluation Relevance and Assessment (MERA) which calls for measurement of all multilaterals against a limited set of generic indicators. Both Denmark's Performance Management Framework (PMF) and the Dutch Multilateral Monitoring System (MMS) employ a measurement process relying on feedback from representatives in the field following specific questions complemented with consultations. The Common Performance Assessment System (COMPAS) by the Multilateral Development Banks (MDBs) is well advanced in the definition of indicators" (CIDA 2006b, vi).

Sources of information for measurement by donors include, in order of importance, joint donor evaluations of multilateral projects or programs, monitoring and evaluation reports produced by the multilaterals themselves, project-specific field reports and the deliberations of Governing Boards. Of lesser importance in this respect are financial audit reports and financial statements. Feedback from the Multilateral Organizations Performance Assessment Network (MOPAN) has become an important source of information for many donors.

There appears to be much agreement about current challenges in the field of performance measurement. Many respondents mentioned the inadequate reporting on the Millennium Development Goals (MDGs). They are keenly aware of the gaps in MO reporting on results in the field. They recognize the need for independent means for evaluating the accuracy and findings of evaluation reports. Without exception there is great interest in harmonization of measurement approaches in line with the Paris Declaration. Measurement of the extent to which programs have been aligned with partner governments is important.

The above challenges indicate the general scope for a strategy to enhance measurement practices. Key components of such a strategy include increased attention to the quality of monitoring and evaluation practices of the multilaterals themselves in order to make them

more results oriented. Reporting by the multilaterals should reflect on progress towards the MDGs, analyze the extent of alignment of programs with the recipient country's own strategies, and indicate application of lessons learned. A well-coordinated process of harmonization of measurement systems among donors would reduce transaction costs and allow donors to form a common front towards change. In addition, some type of institutional arrangement would be useful to assess independently the accuracy of performance measurement in order to reduce the "evaluation gap".

Perhaps the most striking result of the survey was the unanimity in feeling that harmonization of measurement practices is of first priority. It would be in line with the Paris Declaration; it would avoid overburdening organizations with a multitude of donor-specific indicators; it would be an important means of exercising leverage and ensuring greater effectiveness; and, as mentioned, it would reduce transaction costs (Martens et al. 2001; Martens 2005; Ostrom et al. 2005).

3 Achille's heel: managing performance in development agencies

Performance assessment and management can resolve certain problems but also create new problems (cf. Bouckaert / Peters 2002, 359). Having a range of new management practices in place with inadequate or even counterproductive performance assessment and management systems may be worse than having had no reform at all. Implementing an inadequate system of performance management can provide a false sense of security and accomplishment and in the process may misdirect resources and activities. Paradoxically, therefore, inadequate performance management can become the Achilles' heel of development agencies' results agenda itself.

3.1 Four types of performance management approaches

In order to make meaning of the diverse uses and combinations of multilateral performance assessment frameworks, a typology of four models may be used which has been developed by Bouckaert / Halligan (2006, 446 ff.). The model considers three components: (i) performance measurement without management, (ii) management of performance of specific functions and (iii) performance management.

Based on the logical sequence of, firstly, collecting and processing performance data into information; secondly, integrating it into documents, procedures, and stakeholder discourses; and thirdly, using it in a strategy of improving decision-making, results, and accountability, four 'pure' models have been constructed by the two authors (see table 3).

	<i>Model 1:</i> Traditional/Pre- Performance	<i>Model 2:</i> Performance Administration	<i>Model 3:</i> Management of Performances	<i>Model 4:</i> Performance Management
1. Measuring	Intuitive	Administrative data registration	Specialised performance measurement systems	Hierarchical performance measurement system
2. Integrating	None	Some	Within different systems for specific management functions	Systematically integrated
3. Using	None	Limited	Disconnected	Coherent, integrated, comprehensive, consistent
4. Limitations	Functional unawareness	Ad hoc selective	Incoherence	Sustaining complex system
Source: Bouckaert / Halligan (2006, 446)				

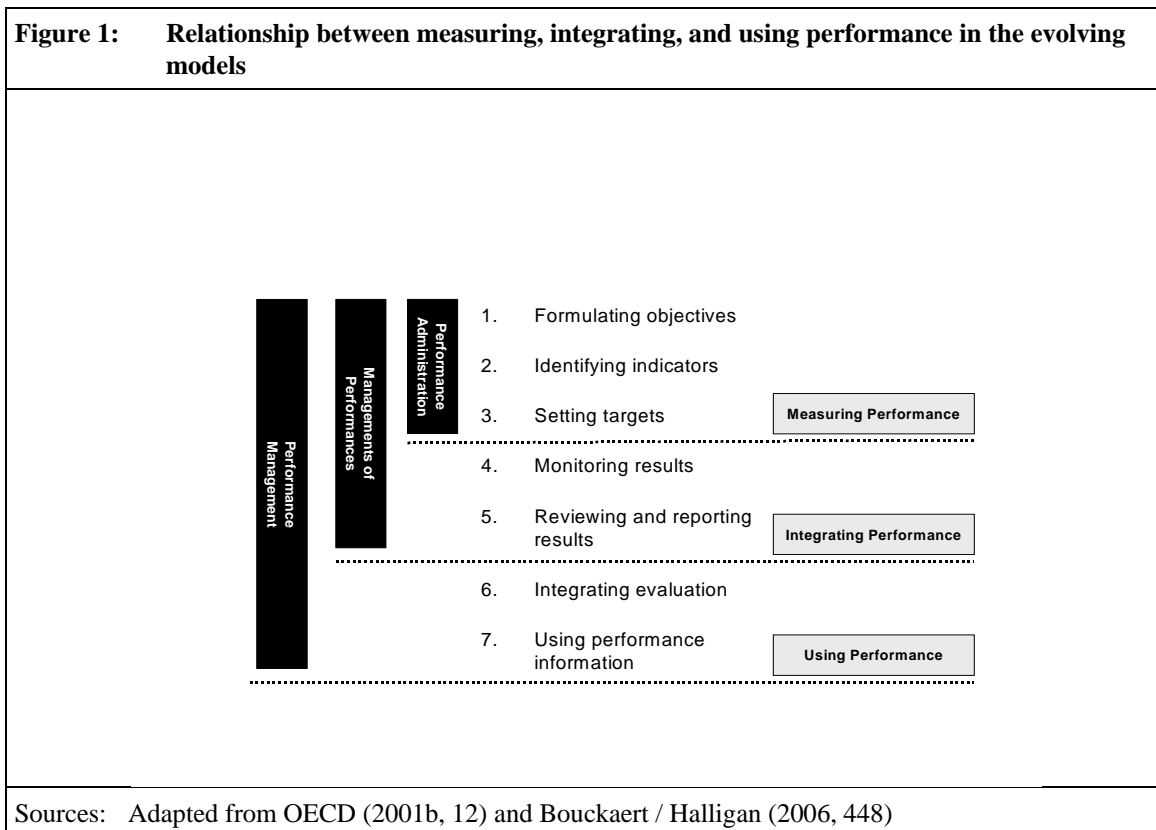
How DAC members' or individual development agencies' institutional differences impact on shaping multilateral performance policies and practice is an area which has often been ignored. Only ODI's (2005) comparative survey included 'institutional' dimensions, such as differences in the role of parliaments or in the engagement style of bilateral donors. The four models distinguished by Bouckaert and Halligan can be applied to the historical and institutional development of performance and management systems as a basis for comparatively analysing and effectively harmonising DAC members' orientations to multilateral organisation performance assessments.¹⁹

Model 1 is termed *traditional/pre-performance* and essentially recognises that multilateral performance objectives in a generic sense can be found in most systems of development agency, but that many of them might be regarded as pre-modern management. In these cases, the expectation is that multilateral performance would be generalised and diffuse, with goals not defined in terms of performance as such as it arises where assessment and management are present. Input driven development agencies, within law-based systems focusing on procedure and due process, may have a very implicit interest in multilateral performance. Scarcely available data may not have an information value, will not be integrated, and hardly will be usable.

Under model 2, *performance administration*, a commitment to multilateral assessment and management is expected, but the relationship may not be explicit or well developed and the application is often ad hoc. The performance administration model focus on multilateral assessment is inclined to be technically oriented but the level of coherence may depend on which generation of assessment system is under discussion. In respective development agencies is an intuitive and generalised concern for multilateral performance that is registered and administered. Multilateral organisation performance assessment becomes

¹⁹ See also Bouckaert's earlier comparative research and typologies of results-based management systems within OECD (OECD 1997; Beuselinck 2006; Campbell 2004).

another administrative procedure that may be part of an administrative and legal setting, not a managerial or policy context. Multilateral information generated from these administrative procedures is disconnected from improvement strategies. Sophisticated rules developed for registering and administering multilateral performance (e.g. institutional strategies) are not developed to generate information to affect managerial functions nor elements of a policy or program cycle.



Model 3 is a quite complex type, entitled *managements of performances*. It arises where management and performance have been linked but the connection between them is underdeveloped and concurrent systems operate. Managing performances includes multilateral performance assessment but goes beyond its administration. It implies different types of performances according to different and unconnected management functions. This results in a divers range of managements of performances, like performances in partnership management, financial management (budgets, audits), strategic and operational management on country-level.

A diverging set of performance measurement systems is feeding information into a disconnected set of management functions, resulting in different multilateral performances for different purposes, which are not necessarily linked in a hierarchical and logical way. A symmetrical development of these function-based multilateral measurement systems make it not very consistent, coherent, comprehensive, and integrated between these functions. However, within some functions there may be a high level of sophistication and development, even up to driving an improvement and reform process in other functions. Accountability-based performance may drive organisational learning and allocation or vice versa.

The Model 4, *performance management*, is defined by the presence of distinctive features: coherence, integration, consistency, convergence, and comprehensiveness. It includes a solid multilateral performance measurement system beyond administration and proliferation. It includes an integration of performance information, which goes beyond ad hoc connectedness, for the purpose of using it in a coherent management improvement strategy. It considers different performance assessment systems for different purposes, which are (hierarchically) connected. The performance management model also requires an explicit policy on assessment and managing the different functions and their performances.

3.2 Donor readiness assessment for multilateral performance assessment

The last 5 years saw the mushrooming of multilateral performance assessment frameworks. Current emphasis by development agencies on joined-up approaches and harmonisation signal a period of performance expansion and management consolidation. A purpose of the four ideal types is to use them as a basis for examining the variations in the orientation of the development agency systems to multilateral performance management. The importance of the donor countries' and their respective development agencies' cultural and administrative traditions is recognised.

The success of RBM systems depends on the contextual conditions in which they are undertaken and implemented. Kusak and Rist, who work for the World Bank on such systems, suggest asking eight "readiness" questions for what they call "Results-based Management and Evaluation Systems" (Kusek / Rist 2002; Kusek / Rist 2004; Kusek / Rist / White 2005). Their questions inquire about the pressures and advocates calling for RBM, the motives of the RBM champion, the beneficiaries of the system, the capacities to support RBM, the way in which the system will help meet program goals, the way it will deal with negative information and how it will be linked to reaching national goals.

RBM reforms prolong in many OECD countries, mainly in English- and Scandinavian-speaking parts of the world. These countries have substantial experience with introducing a results focus and are either implementing or contemplating a new generation of results reforms. Other OECD member countries, especially those with strong legal traditions such as France and Germany, have only recently embarked on such reforms. In Germany most of these RBM reforms are taking place and are limited to the local level (Kusek / Rist 2005, 304).

	Pre/Proto Performance	Managing Performances
Unsystemic	<i>Traditional:</i> ▪ ----	<i>Managements of Performances:</i> ▪ Denmark ▪ Netherlands
Systemic	<i>Performance Administration:</i> ▪ Germany	<i>Performance Management:</i> ▪ Canada ▪ UK
Source: Adapted from Bouckaert / Halligan (2006, 451)		

Performance administration: Germany

Germany's multilateral ODA usually accounts for about a third of total gross ODA but was higher in 2004 (42 %) due to a large contribution to the World Bank. In an increasingly globalised world, Germany attaches growing importance to multilateral co-operation. The last OECD/DAC Peer Review on Germany (OECD 2006c) states, that in order to better focus on the effectiveness of multilateral organisations, BMZ intends to monitor their progress in implementing the internationally agreed upon development agenda, including the fight against poverty and the MDGs at the field level. The internal reorganisation of BMZ in 2003, which combined multilateral, bilateral and sectoral responsibilities across the ministry, has enhanced the consistency between multilateral and bilateral policy making. The shift towards greater focus on performance assessment has not yet been used to guide levels of financial contributions to multilateral agencies. Germany has traditionally supported the European Community (EC) and IFIs but recognises that UN agencies deserve greater support given the role they should play on global issues. This increased focus on multilateralism has yet to be accompanied by similar levels of strategic thinking and performance monitoring. An immediate implication for the increased multilateral focus is, therefore, the formulation of a clear multilateral strategy, and an organised approach involving Germany's implementing agencies which more systematically tracks the performance of the multilateral institutions (cf. *ibid*).

The 'legal state' or *Rechtsstaat* is a key element in the German administrative system. It is characterized by the classical bureaucratic model with strong emphasis on legality and proper fulfilment of regulatory functions (*Ordnungsaufgaben*). Only few reform initiatives are reported at the federal level (Flynn / Strehl 1996). German changes with regard to new management ideas in government reform (e.g., RBM) are characterized more by administrative tightening-up and modernization. According to Pollitt / Bouckaert (2004, 257) the German trajectory has been marked rather by incrementalism ('permanent flexibility of institutional frameworks') than by fundamental change (Benz / Götz 1996, 5; Knill 2001), and more by improvement of the existing system rather than an import of other systems (König 1997).

Accordingly, the DAC Peer Review criticizes greater emphasis on the performance of multilateral organisations could be better translated into a funding allocation policy. The review recommends a more adequate framework for multilateral co-operation, encompassing global funds, based on a methodology that includes an assessment of effectiveness, preferably in collaboration with other donors. Preliminary analysis indicates that Germany currently *administers performance*. So far BMZ works with limited commitment to, and application of multilateral performance, assessment and management principles.

Managements of performances: Denmark and Netherlands

The Netherlands is a strong supporter of the multilateral system. It is the biggest contributor to several UN agencies and has increased its contribution to IDA. It is also committed to increase its focus on European co-operation and is supportive of efforts towards a more unified aid policy among European Union Member States. The multilateral share of Dutch ODA was 28 % in 2005 but approximately four to five percentage points higher if bilateral funds channelled through multilateral agencies in the form of earmarked contributions are taken into account.

The Netherlands is a unitary, but decentralized state (Kickert / In't Veld 1997). The political system is consociational, consensual, multiparty, and corporatist (Lijphart 1999). In the early 1990s, results-oriented budgeting and management regained attention. There was also an increased trend to integrate performance measures in the budget documents. Finally, in 2001, performance-budgeting (VBTB) was legally implemented: the format of the budget bill became outcome oriented and policy objectives and performance measures were integrated in the explanatory memorandum and the accruals-budgeting system has been planned to be extended from the agencies to the departments.

The 2003 policy framework marks a shift in the Dutch multilateral policy towards an approach “more results driven and based on relevance to Dutch policy objectives” (OECD 2006b). This has resulted in a shift from un-earmarked contributions towards theme-based earmarked contributions with the intention of meeting targets in Dutch priority areas. This approach poses some challenges, according to the 2006 DAC Peer Review. First, the resulting fragmentation in Dutch funding of multilateral agencies can make it difficult for the MFA to have an overall consistent approach to multilateral assistance. Second, this *à la carte* practice could hinder multilateral agencies’ management improvement efforts and corporate focus, as individual agencies have to deal with competing demands from members.

In order to increase the overall consistency of its approach, the MFA has been working on a new strategy for managing its multilateral aid portfolio. The key proposal emerging from this work is to focus on about a dozen multilateral agencies and to co-ordinate Dutch contributions to these agencies (at this point this approach has yet to be endorsed). The DAC Review recommends: in reviewing its overall multilateral approach, the Netherlands should make efforts to strike a balance between bilateral considerations and implications for multilateral agencies in order to avoid the risk of “bilateralising” multilateral agency programmes. It should also reinforce co-operation with other DAC members on collective approaches to improving the effectiveness of the multilateral aid system to ensure complementarity among various multilateral assessment initiatives. Finally, given its objective of promoting greater multilateral effectiveness, the Netherlands is encouraged by the DAC Review to elaborate its multilateral strategy and to strike a balance between bilateral objectives and implications for the management of multilateral agencies, taking account of on-going reform efforts (cf. *ibid.*).

The 2003 DAC Peer Review welcomed Denmark's efforts to form coalitions with other donors on issues concerning the performance of multilateral institutions, so as to better co-ordinate the perspectives of all donors in this context. At the level of the field, the “active multilateralism” approach has been recommended being used creatively to promote improved linkages between bilateral and multilateral agencies across all donors. The development of the PMF is part of a larger internal reform of the management of Danish development cooperation, which also includes far-reaching decentralisation of bilateral and multilateral development cooperation. By January 2005, a restructuring of the multilateral cooperation took place, when responsibility for the cooperation with a number of key multilateral organisations was transferred to the four multilateral representations in New York, Washington, Geneva, and Rome. A series of tools and methodologies have been developed to measure, assess, monitor and improve performance and accountability. While the tools for multilateral development cooperation are expected to be further developed, those

used for bilateral cooperation will be further consolidated during the coming years (Danida 2005b).

While multilateral performance management principles and practice (e.g. MOPAN, PMF, MMS, “New Approach”) is fairly well developed, they are inclined to be neither integrated nor systemic yet. Even though, Danida is undertaking major steps in the direction of integrating its *performances managements* (Danida 2005b).

Performance management: Canada and UK

Canada's participation in multilateral programmes is part of its foreign policy agenda of economic prosperity, domestic security, and the projection of Canadian values. Canada works with multilateral institutions because it considers them to be the best way of addressing key development challenges and providing global public goods. The overall objectives with the multilateral institutions are to strengthen their ability to advance human development, particularly, in meeting the MDGs (OECD 2002, 30).

Canada is a federal state with a ‘Westminster’ system. Antibureaucratic, pro-private-sector reforms, borrowed at least in part from the USA and the UK, have been introduced in the 1980s. End of the 1990s however, a ‘Canadian model’ of public sector reform has been proclaimed. This included a rejection of the proposal that minimizing government was always a good thing and an embracing of experiment and diversity in organisational forms (Bourgon 1998; Savoie 1997). A perceived dissatisfaction with government, and alleged citizen demands for greater accountability, were seen as a major factor for public management reform in that time (Pollitt / Bouckaert 2004, 225). Bourgault and Carroll (1997, 97) furthermore, refer to a quite strong ‘mandarin culture’, with considerable horizontal communication between senior civil servants in different departments.

Major RBM-related reform packages include: in June 1995, a Quality Services Initiative approved by the Cabinet, aimed at increasing measured client satisfaction; in 1996, introduction of the Improved Reporting to Parliament Project (IRPP) with the aim of enhancing accountability of ministers and departments to Parliament, in 1997, publication of Accounting for Results (President of the Treasury Board 1997), which for the first time brought together results statements for all departments. The implementation process of public management reforms is in general seen as a somewhat uncertain one, at least in the sense that a number of reforms petered out after a relatively short period, leaving only traces rather than significant achievements that had been predicted at the time of their initiations. It is referred to a significant implementation gap (cf. Auditor General of Canada 1993; cf. Pollitt 2005, 228).

CIDA has been one of the pioneers in results-based management (RBM) and supports the current international effort, centred around the MDGs, to apply RBM across the whole aid system, multilateral as well as bilateral (OECD 2002). Already in 1995, CIDA reported that it had focussed its initial efforts on organizing and planning for renewal mainly at its headquarters. The Agency had developed a results-based management concept to strengthen its effectiveness. CIDA has introduced a Policy on Results-Based Management and is implementing it. The Agency has since developed a Framework of Results and Key Success Factors for measuring its results at the project level. It has also identified a set of Assessment Tools to enable the Framework to be used by operating personnel. These are the first of several steps required to build an effective performance measurement system.

They begin to respond to CIDA's concern that it did not have satisfactory procedures to measure and report on its effectiveness (CIDA 1998; 2002; 2003).

In his November Report 2006 the Auditor General of Canada has urged CIDA to move ahead with determination as it makes the changes necessary to renew itself. He stressed that challenges remain, namely: the need to do a better job of explaining how CIDA's projects contribute to government priorities; the need to establish clear accountabilities not only within CIDA but also with its partners; and the need for more transparency in its dealings with stakeholders and the public. This was a follow-up of the 1993 audit, which had questioned the value achieved with CIDA's expenditures and had made recommendations to improve the Agency's accountability and management effectiveness.

With about one-third of its ODA disbursed to multilateral organisations, the United Kingdom is an important contributor to that system. DFID has developed a strategic approach to the multilaterals, with individual institutional strategies framing the shape of each partnership. DFID tracks the results of the multilateral institutions through a Multilateral Effectiveness Framework (MEFF). These tools help DFID prioritise its multilateral support and provide the means to allow it to play a leading role within the international community to promote the effectiveness of multilateral aid, in line with its 2005–08 PSA objective. The DAC commends the UK as being currently seen by many aid practitioners and donors as one of the bilateral models for today's evolving world of development co-operation (cf. OECD 2006d).

The UK is a unitary and highly centralized state. In 1999 the Prime Minister issued a white Paper, *Modernizing Government*. Subsequently, the increases in public spending were accompanied by further intensification of central target-setting and performance measurement (Hood 2002). As Blair famously said, his second term came to be about 'delivery, delivery, delivery' (cf. Pollitt 2005, 295).

In line with a new approach across the UK Government, DFID set itself the task of basing policy on evidence, focusing on outcomes rather than inputs, and increasing the transparency of policy making and use of resources. The 1997 White Paper shifted the department's measurement of its performance towards the UK's contribution to meeting the Millennium Development Goals, rather than on input-based measures of the UK's contribution. An example of the effort to increase leverage was the Government's willingness to channel resources through multilateral institutions where they are effective, such as through the World Bank. The UK gave a higher share of aid as multilateral assistance than the DAC average. Over 1996–2000 DFID provided on average 41 % of its aid through multilateral organizations compared with 36 % for all DAC donors (Barder 2005, 25 ff.). However, according to (Barder 2005, 26) and in common with other organizations, DFID did not satisfactorily resolve the tension, on the one hand wanting to increase the importance of measuring outcomes, and on the other hand needing evidence to attribute those outcomes to the UK's own contribution.

The 2006 DAC Peer Review (OECD 2006d) suggests that the UK should seek to improve strategic tools for the assessment of multilateral performance, such as the MEFF, and to further maximise their use internally and internationally. It is further recommended, while developing a strategic vision for funding of core and non-core multilateral budgets, DFID should take care not to distort multilateral principles.

For the *performance management* type, Canada and the UK come closest to the stringent requirements, although there are weaknesses in both cases that make the fit less than optimal.

4 Some conclusions and remaining challenges

The making and harmonisation of multilateral performance assessments is now intensified by several donors. In most cases it is still work in progress at best. In other cases scepticism prevails about a lack of evidence regarding their usefulness. Hence it is still not clear to what extent multilateral performance measures help bilateral agencies achieve the goals they have been set by their policy makers or as have been put forward by the international community in the MDGs and the Paris Declaration (Halachmi 2005; Propper / Wilson 2003; Swiss 2005).

Management initiatives from managing by objectives to results-based management often suffer from their own enthusiasm. Expectation gaps are quite common not only for multilateral organisation performance assessment but also more broadly for any project of this type. Ammons / Coe / Lombardo (2001, 104 f.) amplify that even high-quality assessment systems delivering good results are disappointing when expectations are overblown and observers anticipate fabulous results. The battlefield of public management reforms is therefore said to be littered with casualties of excessive hype.

Ammons' et al. analysis of three different 'performance-comparison projects' of local government suggests differences in the gap of initial expectations and perceived results of participants, depending on major objectives and practices²⁰. For example, expectations were high for the identification of better performance measures for agencies' services. These expectations were met fairly well. Moderate success was also recorded for providing unit-cost comparisons, identifying top performers, fostering communication about processes and outcomes within and among participants, and providing information useful for communication with (superior) executive or legislative bodies. Only one out of the three projects was able to provide useful information for operating or budgetary decisions. While none was awarded a critical score for providing information useful in communicating with the general public or for identifying and analysing best practice.

Bevan / Hood (2004) point out that complete transparency in setting targets leads to problems in securing effective control; or in more relative terms: too much predictability in the name of transparency and accountability weakens control (Pollitt / Bouckaert 2004, 159–181). Yet, it is argued that the obvious difficulties do not mean that targets and RBM should be scrapped. Targets have been introduced because alternative forms of control and accountability have been tried and found defective. As Winston Churchill said of democracy, regulating performance through targets is the worst control system ever devised – except for all others (cf. *ibid*, 598). What is needed are ways of limiting gaming.²¹

20 See table 1: "taxonomy of multilateral performance assessments" or "factors affecting the choice of methodology" in OECD / World Bank (2007).

21 See p. 4: the dilemmas of 'creaming', 'teaching to the test' or 'setting low goals' in RBM systems.

One way of doing so is to replace enthusiasm by practicality and to reduce complexity by identifying ‘minimum conditions’ (OECD / World Bank 2007, 3 ff.) for credible and quality multilateral organisation performance assessments ‘under real-world constraints’ (World Bank 2006b, 2 f.). In that respect the last 3 of the 5 MfDR principles (OECD 2005b; 16) are especially relevant, e.g.: (i) keeping results measurement and reporting as simple, cost-effective, and user-friendly as possible; (ii) managing for, not by, results, by arranging resources to achieve outcomes, and (iii) using results information for learning and decision making as well as reporting and accountability.

What else? Multilateral organisation performance assessment requires ‘scaling up’ of conventional RBM approaches applied at project, country or agency levels (Duflo 2005; Ferroni / Mody 2002) and last but not least ‘opening the black box of aid effectiveness’ (Bourguignon / Sundberg 2007). Bourguignon and Sundberg reveal that analyses on aid effectiveness have yielded unclear and ambiguous results. This is not surprising given the heterogeneity of aid motives, the limitations of the tools of analysis, and the complex causality chain linking external aid to final outcomes. The causality chain has been largely ignored and as a consequence the relationship between aid and development has been mostly handled as a kind of black box. Making further progress on aid effectiveness requires opening that box and identifying ‘causal clusters’ (Young 2002) of aid, effectiveness and results, i.e. the often conjured up attribution gap. Multilateral evidence on aid effectiveness is particularly fragile compared to other levels of assessment. Problems of coordination, (lack of) ownership, and attribution make not only financing of (Birdsall 2004) but also accountability for multilateral policies costlier for donors to arrange, and riskier in terms of their sustainability and benefits.

Dealing with both the ‘expectation gap’ and the ‘attribution gap’ remain major challenges for multilateral organisation performance assessments by donors.

Bibliography

- Ammons, D. N. / C. Coe / M. Lombardo* (2001): Performance-Comparison Projects in Local Government: Participants Perspectives, in: *Public Administration Review* 61 (1), 100–10
- Auditor General of Canada* (1993): 1993 Report. Chapter 6: Canada's Public Service Reform, and Lessons Learned from Selective Jurisdictions, Ottawa
- Barder, O.* (2005): *Reforming Development Assistance: Lessons from the UK Experience*, Washington, DC: Center for Global Development
- Behn, R. D.* (2003): Why Measure Performance? Different Purposes Require Different Measures, in: *Public Administration Review* 63 (5), 586–606
- Benz, A. / K. A. Götz* (1996): The German Public Sector: National Priorities and International Reform Agenda, in: A. Benz / K. A. Götz (eds), *A New German Public Sector? Reform, Adaptation and Stability*, Aldershot: Dartmouth, 1–26
- Beuselinck, E.* (2006): *Institutional and Organisational Arrangements of Donor's Aid Administration: An Analytical Perspective on Shifting Coordination Initiatives: Towards an Assessment of the Relevance of a Neo-Institutional Framework*, Milan: European Group of Public Administration
- Bevan, G. / C. Hood* (2004): Targets, Inspections, and Transparency, in: *British Medical Journal* 328 (7440), 598
- Bird, S. M. et al.* (2005): Performance Indicators: Good, Bad, and Ugly, in: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168 (1), 1–27
- Birdsall, N.* (2004): *Underfunded Regionalism in the Developing World*, Washington, DC: Center for Global Development
- Bouckaert, G. / B. G. Peters* (2002): Performance Measurement and Management: The Achilles Heel in Administrative Modernization, in: *Public Performance & Management Review* 25 (4), 359–62
- Bouckaert, G. / W. van Dooren* (2003): Performance Measurement and Management in Public Sector Organizations, in: T. Bovaird / E. Loeffler (eds), *Public Management and Governance*, New York and London: Routledge
- Bouckaert, G. / J. Halligan* (2006): Performance and Performance Management, in: B. G. Peters / J. Pierre (eds), *Handbook of Public Policy*, Thousand Oaks, Calif.: Sage, 443–60
- Bouckaert, G. et al.* (2006): Results-Based Management: Opportunities for Effectiveness and Accountability, in Andreas Obser (ed.), *Distributed Governance Learning and Innovation*, a joint pilot project of the Central Institute for Economic Management (CIEM), Hanoi, Vietnam and the University of Potsdam, Potsdam, Germany, with support from the German Federal Ministry for Economic Cooperation and Development (BMZ) and German Agency for Technical Cooperation (GTZ), Potsdam: University of Potsdam (Potsdam Issue Brief, PIB 2)
- Bourgault, J. / B. Carroll* (1997): The Canadian Senior Public service: The Last Vestiges of the Whitehall Model?, in: J. Bourgault / M. Demers / C. Williams (eds), *Public Administration and Public Management Experiences in Canada*, Quebec: Les Publications du Québec, 91–100
- Bourgon, J.* (1998): *Fifth Annual Report to the Prime Minister on the Public Service of Canada*, Ottawa: Privy Council Office
- Bourguignon, F. / M. Sundberg* (2007): *Aid Effectiveness: Opening the Black Box*, paper given at American Economics Association annual meeting, Chicago, IL, January 6
- Campbell, J. L.* (2004): *Institutional Change and Globalization*, Princeton, NJ: Princeton University Press
- Chancellor of the Exchequer* (2004): *Modern Public Services for Britain: Modernisation, Reform, Accountability: Public Service Agreements 2005–2008*, London: The Stationary Office
- CIDA* (1998): *Lessons Learned from Implementing Results-Based Management (RBM) at CIDA*, Ottawa
- (2002): *Review of Current RBM and Accountability Practices in CIDA*, Gatineau
- (2003): *Results-Based Management: Towards a Common Understanding Among Development Cooperation Agencies*, Gatineau (Discussion Paper)
- (2006a): *Criteria for Positioning Multilateral Organizations Based on the Multilateral Effectiveness and Relevance Assessment (MERA)*, Gatineau

- (2006b): *Bilateral Methodologies for Assessing Multilateral Performance: Survey Findings*, Gatineau
- (2006c): *Interim MDIs Relevance and Effectiveness Assessment (MERA)*, Gatineau
- Danida* (1991): *Effectiveness of Multilateral Agencies at Country Level: Case Study of 11 Agencies in Kenya, Nepal, Sudan and Thailand*, Copenhagen: Ministry of Foreign Affairs of Denmark
- (2005a): *A Results-Oriented Approach to Capacity Development and Change*, Copenhagen: Ministry of Foreign Affairs of Denmark
- (2005b): *Performance Management Framework for Danish Development Cooperation 2006–2007*, Copenhagen: Ministry of Foreign Affairs of Denmark
- (2006a): *Peer Assessment of Evaluation in Multilateral Organisations: United Nations Development Programme*, Copenhagen: Ministry of Foreign Affairs of Denmark
- (2006b): *Peer Assessment of Evaluation in Multilateral Organisations. Peer Review of UNDP Evaluation Office. Key Issues and Lessons Learned*, Copenhagen: Ministry of Foreign Affairs of Denmark
- DFID* (2005a): *The MEF Methodology: A Review of DFIDs Multilateral Effectiveness Framework*, London
- (2005b): *A DFID-Wide Approach to Getting the Best Out of Multilateral Partners*, London
- (2006): *Canada, Sweden and the UK: A Joint Institutional Approach, Working together with UNICEF for the Worlds Children*, London
- Duflo, E.* (2005): *Scaling Up and Evaluation*, paper given at Annual World Bank Conference on Development Economics (ABCDE), Amsterdam
- Ferroni, M. / A. Mody* (eds) (2002): *International Public Goods: Incentives, Measurement, and Financing* Boston and Washington, DC: Kluwer Academic Publishers and the World Bank
- Flynn, N / F. Strehl* (eds) (1996): *Public Sector Management in Europe*, London: Prentice Hall/Harvester Wheatsheaf
- George, S.* (1992): *The Baldrige Quality System*, New York, NY: Wiley
- Halachmi, A.* (2005): *Performance Measurement: Test the Water Before You Dive In*, in: *International Review of Administrative Sciences* 71, 255–66
- Heinrich, C. J.* (2003): *Measuring Public Sector Performance and Effectiveness*, in: B. G. Peters / J. Pierre (eds), *Handbook of Public Administration*, London: Sage
- Hood, C.* (2002): *Control, Bargains, and Cheating: The Politics of Public-Service Reform*, in: *Journal of Public Administration Research and Theory* 12 (3), 309–32
- Hood, C et al.* (1999): *Regulation Inside Government: Waste Watchers, Quality Police, and Sleaze-Busters*, Oxford: Oxford University Press
- Kaul, I.* (2004): *Global Public Goods: A Key to Achieving the Millennium Development Goals*, *The Courier ACP-EU* 202, 15–18
- Kickert, W. J. M. / R. In't Veld* (1997): *National Government, Governance and Administration*, in: W. J. M. Kickert / F. van Vught (eds), *Public Policy and Administration Sciences in the Netherlands*, London: Prentice-Hall/Harvester Wheatsheaf, 45–62
- Knill, C.* (2001): *Institutionelle Indikatoren für die Entwicklungsdynamik nationaler Verwaltungen: administrative Reformkapazität in föderalen und unitarischen System*, paper given at Gemeinsame Tagung der Deutschen Vereinigung für Politische Wissenschaft (DVPW), der Österreichischen Gesellschaft für Politikwissenschaft (ÖGPW) und der Schweizerischen Vereinigung für Politische Wissenschaft (SVPW) zum Thema: *Der Wandel föderativer Strukturen*, Berlin, Juni 8–9
- König, K.* (ed.) (1997): *Entrepreneurial Management or Executive Administration: The Perspective of Classical Administration*, in: W. J. M. Kickert (ed.), *Public Management and Administrative Reform in Western Europe*, Cheltenham: Edward Elgar, 213–32
- Kusek, J. Z. / R. C. Rist* (2005): *Performance-Based Monitoring*, in: S. Mathison (ed.), *Encyclopedia of Evaluation*, Thousand Oaks, Calif.: Sage, 300–05
- / – (2004): *Ten Steps to a Results-Based Monitoring and Evaluation System*, Washington, DC: World Bank, 4
- / – (2002): *Building Results-Based Monitoring and Evaluation Systems: Assessing Developing Countries Readiness*, in: *Zeitschrift für Evaluation* 1, 151–58

- Kusek, J. Z., R. C. Rist / E. M. White* (2005): How Will We Know the Millennium Development Goal Results When We See Them? Building a Results-based Monitoring and Evaluation System to Give Us the Answers, in: *Evaluation* 11 (1), 7–26
- Lijphart, A.* (1999): *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*, New Haven, Conn.: Yale University Press
- Martens, B.* (2005): Why Do Aid Agencies Exist?, in: *Development Policy Review* 23 (6), 643–63
- Martens, B. et al.* (eds.) (2001): *The Institutional Economics of Foreign Aid*, Cambridge: Cambridge University Press
- MfDR* (Managers for Development Results) (2006): Workshop Report – Managing for Development Results: Uganda, June 15–16, 2006, Paris
- MOPAN* (Multilateral Organizations Performance Assessment Network) (2006): *The MOPAN Survey 2006: Perceptions of Multilateral Partnerships at Country Level*, Bern: DEZA
- Multilateral Development Banks Working Group on Managing for Development Results* (2006): 2005 Report: MDB Common Performance Assessment System (COMPAS), Manila: ADB
- Nicholson-Crotty, S. / N. A. Theobald / J. Nicholson-Crotty* (2006): Disparate Measures: Public Managers and Performance-Measurement Strategies, in: *Public Administration Review*, 66 (1), 101–13
- Overseas Development Institute* (ODI) (2005): *Relationships of other Donor Organisations with Multilaterals*, London
- OECD* (Organisation for Economic Cooperation and Development) (1997): *In Search of Results. Performance Management Practices*, Paris
- (2001a): *Results Based Management in the Development Co-operation Agencies: A Review of Experience. Background Report*, Paris
 - (2001b): *Results Based Management in the Development Co-operation Agencies: A Review of Experience. Executive Summary*, Paris
 - (2002): *DAC Peer Review: Canada*, Paris
 - (2005a): *Managing Aid: Practices of DAC Member Countries (DAC Guidelines and Reference Series)* Paris
 - (2005b): *Managing for Development Results. Principles in Action: Sourcebook for Emerging Good Practices*, Paris
 - *The 2006 Survey on Monitoring the Paris Declaration: updated 2 May 2006*, Paris online: http://www.oecd.org/document/52/0,2340,en_2649_15577209_36162932_1_1_1_1,00.html (accessed 24 May 2006)
 - (2006b): *DAC Peer Review: Netherlands*, Paris
 - (2006c): *DAC Peer Review: Germany*, Paris
 - (2006d): *DAC Peer Review: United Kingdom*, Paris
- OECD* (Organisation for Economic Cooperation and Development) / *World Bank* (2006): *Managing for Development Results: Principles in Action: Sourcebook for Emerging Good Practice*, Paris, First Issue, 134
- (2007): *Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards* Paris, Washington, DC
- Ostrom, E. et al.* (2005): *The Samaritan's Dilemma: The Political Economy of Development Aid*, New York, NY: Oxford University Press
- Pollitt, C.* (2005): *Performance Management in Practice: A Comparative Study of Executive Agencies*, in: *Journal of Public Administration Research and Theory* 16 (1), 25–44
- Pollitt, C. / G. Bouckaert* (2004): *Public Management Reform: A Comparative Analysis*, 2nd edn., Oxford: Oxford University Press
- Propper, C. / D. Wilson* (2003): *The Use and Usefulness of Performance Measures in the Public Sector*, Bristol: University of Bristol (CMPO Working Paper Series 03/073)
- Savoie, D.* (1997): *Central Agencies: A Government of Canada Perspective*, in: J. Bourgault / M. Demers / C. Williams (eds), *Public Administration and Public Management Experiences in Canada*, Quebec: Les Publications du Québec, 59–69

- SDC* (Swiss Agency for Development and Cooperation) / *SECO* (State Secretariat for Economic Affairs) (2005): *Switzerland's Multilateral Development Cooperation Strategy: An SDC-SECO Guideline*, Bern
- Swiss, J. E.* (2005): A Framework for Assessing Incentives in Results-Based Management, in: *Public Administration Review* 65 (5), 592–602
- Talbot, C.* (2005): Performance Management, in: E. Ferlie, L. E. Lynn / C. Pollitt (eds), *The Oxford Handbook of Public Management*, Oxford; New York: Oxford University Press, 491–517
- (2006): Performance Regimes and Institutional Contexts: Comparing Japan, UK and USA, paper given at International Symposium on Policy Evaluation, Tokyo, 24–25 June
- UNDP* (United Nations Development Programme) (2006): *UNDP Management Response*, New York
- UNICEF* (United Nations International Children's Emergency Fund) (2006a): *Peer Review of Evaluation Function at United Nations Children's Fund (UNICEF): Final Report*, New York
- (2006b): *Peer Review of Evaluation Function at United Nations Children's Fund (UNICEF): Management Response*, New York
- UNIDO* (United Nations Industrial Development Organization) (2005): *Assessment of UNIDO by Department for International Development (DfID) of the United Kingdom: Summary of Salient Points by the Secretariat*, Vienna
- World Bank* (2004): *Addressing the Challenges of Globalization: An Independent Evaluation of the World Bank's Approach to Global Programs*, Washington, DC
- (2006a): *Annual Review of Development Effectiveness 2006 – Getting Results*, Washington, DC
- (2006b): *Conducting Quality Impact Evaluations under Budget, Time and Data Constraints*, Washington, DC
- (2006c): *Global Monitoring Report 2006: Strengthening Mutual Accountability – Aid, Trade & Governance*, Washington, DC
- Young, O.* (2002): Are Institutions Intervening Variables or Basic Causal Forces? Causal Clusters vs. Causal Chains in International Society, in: M. Brecher / F. Harvey (eds), *Reflections on International Studies at the Dawn of the New Millennium*, Ann Arbor, Mich.: University of Michigan Press

Annexes

Annex 1: Synthesis of international comparison

	MOPAN	COMPAS/ MfDR	MEFF	MERA	PMF	MMS	„New Approach“
Responsible organization(s)	Utstein plus (2003)	Multilateral Development Banks (2005)	DFID/UK (2003)	CIDA/Canada	Danida/Denmark (2004)	MFA/Netherlands	Danida & OECD/ DAC (2004)
Approach	“Perception Analysis”	“Self Assessment” & harmonization of practices	“Organizational effectiveness of MO”)	“Results & Management Assessment”	“Quality assessment of MO”	“Perception analysis”	“Peer Assessment”
Objective	Surveying partnership behaviours of MO with national governments, civil society and other bilateral and other relevant MO at the country level.	(a) pool information about how MDBs are contributing to development results (outputs and outcomes); (b) Monitor and synthesize MDBs’ progress over time; and (c) contribute to lessons-learning, accountability and transparency.	Provide information for agency’s: (a) Public Service Agreement reporting, (b) its Institutional Strategies, and (c) its financing strategies	Guiding CIDA staff in gathering information following a standard template of questions; Enabling responsible officers to make a final judgement on relevance and effectiveness;	(a) enhance quality of Danish development cooperation through focus on results, (b) improve management and continuous learning, through better information and reporting, and (c) strengthen accountability through performance assessments and measurement in the context of increasingly decentralized management structures	Giving an insight into the general perception of a given MO	(a) strengthen evaluation capacity of MO and by that (b) enabling bilateral donors to depend on multilaterals’ own results reporting for performance assessments
Criteria	Focus on quality of (a) national partnership (e.g. contribution to policy-dialogue, advocacy); (b) inter-agency partnership (e.g. information sharing, interagency coordination.	3 components based on seven pillars: (a) developing country level (1=capacity building); (b) agency level (2=financing, 3=strategies, 4=projects/programs, 5=M&E, 6=learning and incentives); (c) inter-agency level (7=working with other donors and investors on the MfDR agenda)	Eight organizational systems and their focus on (a) internal performance, (b) country level results and (c) partnerships	(a) relevance, (b) effectiveness and (c) improvement measures	(a) corporate level (DANIDA); (b) organisation level (HQ); (c) country level(field level)	Set of qualitative perception questions	(a) credibility of MO’s evaluation process and reports; (b) independence of MO’s evaluation system, (c) use of evaluation evidence by MO;

Methodology/ indicators	(a) standard questionnaires, (b) agency templates, (c) country team discussions, (d) results of group discussions, (e) synthesis report Asking MO for 'management responses' to MOPAN reports to be posted also on the website	Process and results indicators	(a) Checklist, (b) scorecard format and a "traffic lights" scoring system, and (c) summary report Grouping of MO, combined on the basis of functionality; Process indicators	MOs rated on a scale of 1 to 5 for each of the following 3 areas: relevance, results achieved and management. Weighing in the actual development results achieved, while problems of attribution are recognized.	(a) organisation strategies, annually measured through indicators of respective action plans; (b) at least biannually high-level consultations of MFA and MO; (c) perception analyses of general MO performance; (d) joint donor assessment of MO (i.e. MOPAN); (e) perception analyses of MO at country level; (f) assessment of MOs M&E systems (i.e. "New Approach")		
Link to 'Paris Declaration' (country level)	Yes (predominant)	Yes (plus country capacity building for MfDR)	yes	yes	yes	Yes (predominant)	Yes
Link to MDGs (international aid architecture)	No (but on coherence with practice in the Rome and Paris Declarations)	Indirectly, as COMPAS report is used as input to the Global Monitoring Report by WB/IMF		yes	yes		
Transaction Costs			Staff time/travel costs lower as compared to external evaluation				
Sustainability of exercise		Commitment by heads of MDBs;	Internal training and dissemination required to compensate: (a) staff turnover, (b) insufficient resourcing of overhead costs, and (c) lack of MEFF-awareness in other DFID parts.				
Limitations or critique			(a) doubts of MOs about generic applicability; (b) concerns of MOs about proliferation of similar initiatives	Suffers from a lack of detail. It is considered to be an interim system requiring elaboration and specificity			

Annex 2: The MOPAN

1. RATIONALE

Public opinion and government decision-makers are paying increasing attention to the effectiveness of the multilateral organisations (MOs) to whom they provide resources. The need for better information about MO behaviour in developing countries has become ever more compelling. On the other hand, the resources that each donor can devote to gathering relevant knowledge of multilateral performance are limited.

Working together allows MOPAN members to meet this challenge. Drawing on the collective knowledge and experience of their country-level staff, as well as encouraging the latter's involvement in and ownership of the exercise helps MOPAN avoid duplication of work on all sides. Pooling of resources keeps transaction costs at a minimum and makes the assessments cost-effective for the participating MOPAN members.

2. APPROACH

MOPAN carries out regular joint assessments of the work of MOs in a number of countries where members have their own bilateral programmes. As a rolling exercise, most of the major MOs at the country-level will, over time, be covered. Maintaining a standard methodology makes it possible to compare results over time and identify trends.

The assessment is an opinion survey. It draws upon the perceptions of MOPAN member staff about the in-country performance of MOs, relative to their respective mandates. Participants are asked to give their views on those behavioural aspects of MOs' performance where they are likely to be knowledgeable thanks to their direct inter-agency contacts.

The assessment focuses on the quality of the partnership behaviour of the MO's:

Their national partnerships (contribution to policy dialogue, capacity development, advocacy, support to civil society, and alignment to national institutions, policies and administration)

Their inter-agency partnerships (information sharing, inter-agency coordination, harmonisation and general local responsiveness)

3. PROCESS AND INSTRUMENTS

The institutional questionnaire

The institutional questionnaire is designed to help each MOPAN member embassy/country office participating in the assessment to assemble its views about MO performance on a range of partnership issues. It is completed by each embassy/country office prior to the group discussions within the MOPAN country team and provides an input to the compilation of the country report.

The templates on the MOs

To assist the country teams, the MOPAN headquarters group prepares for them a background information brief on the key aspects of each MO (e.g. mandate, structure, and organisation).

The hotline

A hotline is at the disposal of the country teams for advice and support during the actual assessment period. The hotline responds rapidly to their queries about the objectives, the approach, the process as well as the use of the Survey's findings.

The country team discussion

The MOPAN country team meets as the focus group, where individual knowledge and perceptions are pooled and a collective view of the MOs' respective performance emerges. This exercise serves also to encourage the staff's ownership of the exercise.

The country reports

The outcome of the group discussions is condensed into the MOPAN country report. It reflects the collective point of view arrived at by the group on each MO. The country report also contains information about the team's process in reaching its consensus.

The Synthesis Report

The Synthesis Report provides a synthesis of the country reports, based largely on a textual analysis of the reports. The aggregate questionnaire results are also used as an input.

The Survey has been implemented at field level during February to April and the Synthesis Report been compiled during May and June. The report has been adopted by the MOPAN HQ team in October and discussions with the MOs have taken place by the end of 2006.

4. DISCUSSIONS WITH THE ASSESSED MULTILATERAL ORGANISATIONS

Dialogue with the MOs at their headquarters

The Synthesis Report is presented to the MOs concerned at their headquarters. This is an opportunity for a substantive dialogue between the MOPAN headquarters group and the MOs and for mutual learning among partners.

Discussions with the MOs at the country-level

At the country-level, the MOPAN country team shares the country report with the respective MO country offices. A follow-up meeting is held once the Synthesis Report has been issued. Sharing the country report and the Synthesis Report provides an opportunity to increase mutual knowledge and understanding among partners.

5. COMMUNICATIONS

The final version of the annual Synthesis Report is posted on the external websites of each of the participating MOPAN members, together with any comments on the report provided by the assessed MOs. A Frequently Asked Questions and a one page fact sheet are also posted on their websites.

Annex 3: Example of MOPAN: UNICEF

Box 1: UNICEF's partnership performance at a glance: summary of findings
<p>UNICEF's partnership performance at a glance: summary of findings</p> <p>According to the MOPAN country reports, the overall perception of UNICEF at the country level is that of a very knowledgeable, committed and influential multilateral organisation in the specific areas of its mandate. UNICEF is perceived to be a strong voice for children and women, to pursue its clear priorities, and to be determined to show results and deliver on its promises - including in very difficult circumstances. This is balanced with a view that at times UNICEF is not fully sensitive to the local or to the broader development context. MOPAN country teams perceive that getting the job done quickly sometimes seems more important to UNICEF than investing in partnerships, especially with civil society and other development agencies.</p> <p>Summary of how MOPAN country teams perceive the quality of UNICEF's partnership behaviour towards national stakeholders:</p> <ul style="list-style-type: none"> a. positive contribution to policy dialogue, but room for further improvement (page 7); b. mixed performance in terms of capacity development of various national stakeholders; more engaged with public institutions where contribution to capacity development is moderately effective; good quality technical expertise – although not always appropriate for local needs (page 8); c. a strong comparative advantage in advocacy (page 10); d. various experiences of involving NGOs in policy dialogue; very little such experience with the private sector (page 12); and e. support to national poverty reduction strategies; progress in aligning its own work; weak alignment with government modalities and procedures with the exception of sector-wide approach arrangements (page 12). <p>Summary of how MOPAN country teams perceive the quality of UNICEF's partnership behaviour towards other development agencies:</p> <ul style="list-style-type: none"> f. rather modest performance in terms of sharing information (page 14); g. active and regular participation in local donor coordination groups; mixed picture regarding coordination at the project/programme level; good examples of UN coordinated responses to crisis situations (page 15); h. slow adapter to local aid harmonisation efforts (page 17); and i. improved responsiveness to local concerns, particularly as a result of increased decentralisation of decision-making authority to country offices and positive contributions of heads of UNICEF country offices (page 18).
Source: (MOPAN 2006, 4)

Annex 4: Managing for Development Results (MfDR) Principles

Box 2: Examples of MfDR in development agencies		
MfDR Principles	Examples of tools being used to manage for results in development agencies	Why these are important
Focus the dialogue on results at all phases of the development process	Agency policy/priority frameworks Country programming strategies Thematic/sector strategies Project results frameworks Stakeholder/partner planning and consultation mechanisms Logic models or results chains (integrated in all of the above)	Results-based tools and processes are used to plan for and implement intermediate results linked to country outcomes. Different tools are adapted depending on the level at which they are used, but they all show how policy, country program, thematic and project results contribute to country, regional, or global outcomes, including supporting the Millennium Development Goals (MDGs).
Align programming, monitoring and evaluation with results	Agencywide multiyear strategic plans or rolling work plans Annual program/project management plans, work plans and budgets Training and guidelines for project/program planning and results management Performance management plans	Results-based tools are used to demonstrate how agency investments and inputs will contribute to country, regional, or global outcomes, as well as to indicate how different agency management processes can support the achievement of results.
Keep results measurement and reporting as simple, cost-effective, and user-friendly as possible	M&E systems, plans and guidelines (incorporating MIS) Audit and risk management frameworks Performance measurement frameworks Program/project monitoring frameworks Audit guidelines and tools Evaluation guidelines and tools Risk analysis guidelines and tools Training and guidelines for indicator design, data collection, and analysis	Tools and guides describe the steps and processes to be used in collecting and analyzing performance data at different levels within development agencies, and form the basis for continuing skill building with agency managers and staff.
Manage for, not by, results, by arranging resources to achieve outcomes	Performance reviews and evaluations Internal/external performance monitoring processes and reviews Performance and management audits Thematic and sector studies	Performance information from monitoring and evaluation is used as the basis for assessing progress toward identified country, regional, or global outcomes at various levels.
Use results information for learning and decision making as well as reporting and accountability	Annual agency performance reports to advisory boards/committees, elected officials and their citizens Annual country program performance reports to agency decision makers/committees, elected officials, and country citizens Training and guidelines for management decision making based on results information Analysis of evaluations and reallocation of resources to better achieve results	Reports within and among agencies, and between agencies and their main stakeholders, provide the basis for ongoing agency strategic review, performance adjustment, and reallocation of resources.
Source: OECD / World Bank (2006, 134)		

Annex 5: The Common Performance Assessment System (COMPAS)²²

In response to international commitments on performance and accountability, a Common Performance Assessment System (COMPAS) was proposed by the Multilateral Development Bank (MDB) Working Group on 'Managing for Development Results' (MfDR). The coalition of MDBs involved in this initiative include: the African Development Bank, Asian Development Bank, European Bank for Reconstruction and Development, Inter-American Development Bank and the World Bank.

COMPAS' objectives are to:

- Pool information about how MDBs are contributing to development results (outputs and outcomes);
- Monitor and synthesize MDBs' progress over time; and
- Contribute to lessons-learning, accountability and transparency.

COMPAS findings have been published in an annual report, in 2006. It is intended to produce an annual report every year. This report will act as an inter-agency communication product, as it will systematically record MDBs' respective MfDR policies and practices, and relevant modifications over time. More importantly, however, it will seek to identify common definitions and results-based methodologies between these organizations that can be integrated into a common framework. COMPAS' goal is to facilitate comparability of data between MDBs (e.g. regarding program outputs and outcomes), and ultimately, to contribute to harmonization of practices between institutions.

Annual updates on MDB's respective progress and lessons learned, and according revisions to the common approach, will allow monitoring of progress. Transparency will be ensured by making the report accessible to all external stakeholders and the public.

The conceptual framework for COMPAS draws on MDBs' existing performance systems, and includes three basic components:

- i) Building MfDR capacity in developing countries by focusing on needs assessment knowledge sharing and capacity building;
- ii) Improving the results focus of their internal systems and instruments in areas of finance, country/sector and regional strategies, projects and programmes, monitoring and evaluation, and learning and incentives; and
- iii) Working with other development agencies (donors and investors) on the MfDR agenda.

COMPAS intends to put forward process and results indicators for subcategories of the above mentioned basic pillars, selecting them on the basis of relevance to results, coverage, comparability and availability of data. The following master matrix of categories and indicators is presented in COMPAS 2005 proposal. The aforementioned three components are based on a total of seven pillars:

²² Source: CIDA (2006b).

Table 5: Components and pillars of COMPAS			
PERFORMANCE CATEGORY	SUB-CATEGORIES TO BE MONITORED	PROCESS INDICATORS	RESULTS INDICATORS
Pillar I: Developing country level			
1. Capacity building (CB)	1a. Public and private sector capacity building needs assessment	- CB needs assessment tools available	- % of countries with needs assessments completed - % country strategies with analysis of country capacity for MfDR
	1b MfDR sensitization and knowledge sharing	- MfDR knowledge sharing mechanisms established (workshops, networks, virtual links etc)	- % target audience in developing countries covered
	1c MfDR capacity building programmes	- CB programmes launched where relevant	- % of target countries with CB programs underway or completed
Pillar II: Agency level			
2. Financing	2a Performance based resource allocation for low-income countries	- country performance indicators in place for eligible countries	- % concessional resources allocated to countries with high scores according to each MDB's PBA formula
3. Strategies (country, sector, regional)	3a Outcome focus, evaluability	- Revised guidelines issued to enhance results focus of strategies - Mechanisms in place for reviewing the results focus of strategies	- % of borrowers with MDB strategies based on revised guidelines - % of country strategies whose evaluability is independently reviewed
4. Projects and programmes	4a Outcome focus, evaluability	- Project results frameworks reviewed and improved where necessary (revised guidelines etc)	- % with satisfactory results frameworks as defined by each MDB - % with satisfactory baseline data - % of projects whose evaluability is independently reviewed
	4b Project and programme delivery (speed of disbursement and implementation)	- Simplification of disbursement procedures - Simplification of operational procedures	- Actual annual disbursement as % amounts available for disbursement at beginning of the year - Ratio of originally planned over active implementation period length
	4c Quality of design and supervision	- Increased arms length scrutiny of project quality at entry - Quality of supervision reviews in place or under development	- QAE % satisfactory or better

	4d Management of project risk	<ul style="list-style-type: none"> - Early warning system in place - Proactive management of projects at risk 	<ul style="list-style-type: none"> - % projects at risk, on alert status or similar - % projects at risk being actively managed (proactivity index)
5. Monitoring and evaluation	5a Monitoring instruments, procedures, practices	<ul style="list-style-type: none"> - Active project monitoring system in place (self-assessed) yielding appropriate levels of information 	<ul style="list-style-type: none"> - Monitoring/supervision compliance rates
	5b Quality assurance systems	<ul style="list-style-type: none"> - Arms length review of self-assessed reporting in place or being developed 	<ul style="list-style-type: none"> - Arms length review reports
	5c Independent evaluation	<ul style="list-style-type: none"> - Ex-post evaluation of country programmes - Ex-post evaluations of projects and programmes 	<ul style="list-style-type: none"> - % active borrowers that have a country programme evaluation - % projects and programmes with satisfactory outcomes
	5d Results reporting system	<ul style="list-style-type: none"> - Adequate resourcing of M&E activities - Regular and extensive reporting of M&E data 	<ul style="list-style-type: none"> - % admin. budget allocated to internal M&E (quality assurance, evaluation and audit) - other indicator to be identified
6. Learning and incentives	6a Capturing and using good practice	<ul style="list-style-type: none"> - Good practice materials collected, field tested and available on-line 	<ul style="list-style-type: none"> - Management uptake of evaluation recommendations as reported to Exec. Boards
	6b Staff training and guidance	<ul style="list-style-type: none"> - MfDR training packages developed and field tested - On-the-job help lines established 	<ul style="list-style-type: none"> - % operational staff participating in MfDR training - other indicator to be identified
	6c Staff incentives	<ul style="list-style-type: none"> - Formal and transparent use of MfDR dimensions in annual personal performance reviews 	<ul style="list-style-type: none"> - % operational staff covered by results- focused incentive system
Pillar III: Inter-Agency level			
7. Working with other donors and investors on the MfDR agenda	7a Information sharing	<ul style="list-style-type: none"> - Participation in MDB MfDR working group - DAC JV on MfDR 	<ul style="list-style-type: none"> - MfDR Sourcebook on line and updated regularly
	7b Harmonisation	<ul style="list-style-type: none"> - Areas for MfDR harmonisation identified by working group 	<ul style="list-style-type: none"> - Number of MfDR products and processes for which common principles are agreed (to be defined)
	7c Cooperative/joint ventures	<ul style="list-style-type: none"> - Procedures developed for joint country, sectoral and regional programming - COMPAS developed 	<ul style="list-style-type: none"> - Number of joint MfDR activities/programmes - COMPAS reporting annually
Source: CIDA (2006b, 29–30)			

Annex 6: The MEFF Checklist and major findings

	Organizational systems	Internal Performance	Focus on country Level results	Focus on partnership
	Corporate Governance	3 questions	3 questions	3 questions
	Corporate Strategy	3 questions	3 questions	3 questions
	Resource Management	4 questions	3 questions	2 questions
	Operational Management	4 questions	2 questions	3 questions
	Quality Assurance	4 questions	3 questions	2 questions
	Staff Management	4 questions	2 questions	3 questions
	M&E Lesson learning	3 questions	3 questions	3 questions
	Reporting	3 questions	3 questions	3 questions

Source: DFID (2005)

Since its inception, MEFF has yielded several general results. It found that most international organizations use RBM, and their approaches are increasingly converging due to improved inter-agency communication and coordination efforts. MEFF has identified three forms of organizational change: *consolidating reforms*, *'big bang' reforms* and *incremental reforms*. Mature agencies that have long had RBM frameworks are those that are consolidating reforms (e.g. World Bank). Formerly weak agencies that adopted RBM systems as part of major structural change are instituting 'big bang' reforms (e.g. UNDP, UNFPA, UNIDO, African Development Bank, etc) . Meanwhile, other agencies that have been slower and more prudent in their adoption of RBM, are making only piecemeal, incremental reforms (e.g. WHO, IFAD, Asian Development Bank).

MEFF found that organizations assessed had strongest results in areas of corporate strategy and human resource management. They also scored high in internal performance, reflecting long-standing concerns with internal efficiency, delivery and quality assurance;

and in inter-agency partnerships, indicating an improvement in inter-agency coordination systems. Conversely, most organizations scored lowest in monitoring and evaluation, focusing more on inputs and activities, than outcomes; and in country level results, due to an inability to bring country programmes in line with national poverty reduction strategies.

MEFF's inter-agency comparison revealed that there are great variations in organizational effectiveness among different groups of multilaterals, most notably among standard-setting agencies. Humanitarian and coordinating agencies scored best overall, and MDBs demonstrated great consistency.

Annex 7: Example of MEFF: UNIDO

Box 3: Example of MEFF: UNIDO

DFID communicated the MEFF outcome to UNIDO on 2 March 2005; permission for public disclosure was received on 31 March 2005. It includes an overview of the results of 23 multilateral institution assessments, information on the MEFF methodology, and a detailed assessment of UNIDO.

Within the UN standard setting agencies, UNIDO was rated the best:

- UNIDO achieved a total score of 86, while the average score was 62 (the other organizations scored 51 to 68 points);
- On internal performance, UNIDO scored 88 points, while the average was 66 (the other organizations scored from 52 to 68 points);
- On the country level results focus, UNIDO scored 75 points, while the average was 51 (the other organizations scored from 28 to 65 points);
- On partnerships, UNIDO scored 95 points, while the average was 67 (the other organizations scored from 54 to 70 points).

Within the 23 multilateral institutions, UNIDO was rated:

- Sixth on total scores;
- Fourth on internal performance;
- Eleventh on the country-level results;
- Second on partnerships.

Out of the 72 indicators used by DFID to assess internal performance, country level results and partnership, 70 indicators were applicable to UNIDO. The following provides a summary of DFID assessment ratings of UNIDO:

Rating	Description	Number achieved	Percentage
	All the assessments are positive	51	73
	There are concerns about some issues, but improvements being put in place – Watching brief	18	26
	The assessments are negative, and no plans for improvement yet	1	1
	No information	0	-
	Not relevant	2	-

Source: UNIDO (2005)

Annex 8: The Multilateral Effectiveness and Relevance Assessment (MERA)

Table 7: Interim MO relevance and effectiveness assessment (MERA) of CIDA			
Source of information	Relevance	Effectiveness	Improvement measures (by the institution and by MPB)
E.g. 2000-2003 - MOPAN 2004 - DANIDA - CIDA Field Staff Survey 2005 - DFID MEFF (Score: 88) - Global Funds Evaluation - CIDA Field Staff Survey	Links to MDGs: <ul style="list-style-type: none"> ▪ General link between MO's mandate and MDGs; ▪ Specific MDGs targeted by MO 	Overall governance <ul style="list-style-type: none"> ▪ General overview of MO; ▪ Donor or recipient driven; ▪ Resident board or executive council ▪ Independent evaluators' impression of governance; ▪ General reputation; ▪ Personal insights of evaluator 	<ul style="list-style-type: none"> ▪ Donor or recipient driven specific improvements (e.g. instituting a report, implementing a new policy, providing consulting to improve RBM) ▪ General improvements (e.g. ensuring strong Canadian voice on board on issues with respect to operating budget harmonization)
	Role in the Multilateral architecture: <ul style="list-style-type: none"> ▪ How MO fits in multi development system Region/ sector specific, UN or IFI or other ▪ Kind of aid loans, grants ▪ Budget support ▪ Tech assistance ▪ Partnership with another MDI ▪ role within the MDI system in a particular area 	Efficient use of resources <ul style="list-style-type: none"> ▪ Operating costs as a percent of its aid delivery ▪ Strengths and weaknesses, factors which hinder efficiency 	
	Support to IPS sectors: <ul style="list-style-type: none"> ▪ Sectoral breakdown of its most recent year of programming 	Donor harmonization and alignment – Paris Declaration <ul style="list-style-type: none"> ▪ Work towards objectives of the Paris declaration ▪ Harmonizing/aligning with other UN organisations or IFIs and bilaterals 	
		Managing for results <ul style="list-style-type: none"> ▪ Extent of RBM systems ▪ Reporting systems 	
		Quality of reporting, monitoring & evaluation systems	
		Gender mainstreaming: <ul style="list-style-type: none"> ▪ Policy ▪ Implementation 	
	Environment mainstreaming: <ul style="list-style-type: none"> ▪ Policy ▪ Implementation 		
Overall assessment: Description of how relevant and effective the institution is both in its own right and compared with other MDIs and donors comes in.			
Source: CIDA (2006c)			

Purpose of Multilateral Effectiveness Review

- (i) To better inform policy and financial allocation decisions on more solid evidence. Ultimately, we aim at getting more predictable multi-year funding, and an adequate balance between core and responsive/earmarked funding.
- (ii) To better exercise our accountability and improve our reporting to Canadian citizens.
- (iii) For more effective Board meetings and better identification of areas requiring improvements.
- (iv) To deliver on MPB commitment to complete the multilateral effectiveness review, as part of the Agency Aid Effectiveness agenda.

As stated in the 2006–07 Report on Plans and Priorities for CIDA:

“The Government of Canada is committed to taking a more strategic approach in its support to multilateral organizations, including promoting reforms for aid effectiveness and transparency, ensuring resources are allocated where they can have the greatest impact, encouraging greater participation by international financial institutions in debt relief, and focusing greater support on those institutions that are most effective in promoting governance (including freedom and democracy, human rights, rule of law, justice, and accountable public institutions) and contributing to the MDGs.

The following provides a grid for comparing MOs receiving core funding within each Directorate. Assessments will be based on the submissions made by Program Managers of the Multilateral Effectiveness and Relevance Assessments (MERAs), supplemented by additional information, such as the institutions reports, the field surveys, any reports indicating results achieved. The next step will be to review the current level of funding to these institutions and propose a new allocation, as required.

The proposed grid is consistent and coherent with the new Agency Logic Model

Methodology

Organizations will be rated on a scale of 1 to 5 for each of the following 3 areas: relevance, results achieved and management.

1	2	3	4	5
disastrous	poor	average	very good	outstanding

Limitations:

BMG decided to weight in the actual development results achieved, although, as noted by DFID in their MEFF Methodology: “There is very little information available on results; there are problems in attributing results to the actions of a single aid agency; it is difficult to compare different types of results (e.g. inoculations, enrolments, policy advice, capacity building)”.

EXPLANATION OF CRITERIA

RELEVANCE

1. MDGs:

- a) To what extent does the MO's mandate, corporate strategy and programming support the MDGs?
- b) Has the MO integrated MDG-like targets and indicators in planning processes, both globally and on a country/regional basis?

2. Other Canadian international development policies:

- a) To what extent does the MO support other Canadian international development policies in general?
- b) Does the MO have a policy on mainstreaming Environment?
- c) Does the MO have a policy on mainstreaming Gender Equality?

3. Multilateral architecture:

- a) How relevant and strategic is the MO within the multilateral development system?
- b) Does the organization play an important international role in setting norms and practices? Does it play a pro-active and leadership role on select sector/thematic issues?
- c) Does the mandate of the organization respond to legally binding obligations?
- d) Does the mandate of the organization overlap with others, or is it unique, such as humanitarian assistance, and has it established a clear division of labour with peer multilateral development institutions?
- e) Does the organization have adequate formal and informal collaborative arrangements/partnerships/linkages with like-minded organizations to support performance and bring new ideas and resources to the organization?

RESULTS

These correspond to the **immediate outcomes** level in CIDA's logic model.

1. Institution's own **performance rating**: how does the MO rate itself on results?
2. Is the MO achieving expected and **sustainable results** on the ground, including capacity building, or timely and effective delivery of humanitarian assistance?
3. **Advocacy role**: what evidence is there of partner countries which have modified their policy frameworks as a direct result of advocacy and influence by the MO (e.g. integration of MDGs in poverty reduction strategies)?

- 4. Setting norms:** “has the MO established or increased partnerships which support knowledge-sharing and policy influence to set or raise standards of development, for example for the treatment of disease or the definition of sound governance?”

MANAGEMENT OF THE INSTITUTION (Efficiency & Effectiveness)

These correspond to the **Immediate Outcomes** in CIDA’s logic model

1. Efficiency:

- a) What has been the trend in overhead costs? Have there been recent exercises to increase the cost-effectiveness of administration, such as operational budget reviews?
- b) Is the organization’s financial and budgetary management effective, accountable and transparent?
- c) Are funding sources secure, stable and diversified?

2. Effectiveness – Governance:

- a) Is there a broadly-based governing board, particularly in terms of representation from developing countries?
- b) Is there a clear **accountability** framework upon which management/staff performance is assessed?
- c) Is there an independent **oversight** function?

3. Effectiveness – Harmonization:

- a) To what extent does the MO support Program Based Approaches, defined as a way of engaging in development cooperation based on principles of co-ordinated support for a locally owned programme of development, such as a national development strategy, a sector programme, a thematic programme. Programme based approaches share the following features:
 - (i) leadership by a host country or organisation;
 - (ii) a single comprehensive programme and budget framework;
 - (iii) a formalised process for donor co-ordination and harmonisation of donor procedures for reporting, budgeting, financial management and procurement;
 - (iv) efforts to increase the use of local systems for programme design and implementation, financial management, monitoring and evaluation.

Other macro indicators as defined by the Paris Declaration include:

- (b) A high proportion of the MO’s aid flows are reported on partners’ national budgets;
- (c) A high proportion of capacity development support from the MO provided through co-ordinated programmes consistent with partners’ national development strategies;

- (d) A high proportion of aid flows from the MO use procurement and/or public financial management systems in partner countries;
- (e) A minimal usage of parallel project implementation units;
- (f) A high predictability of aid disbursements from the MO released according to agreed schedules in annual or multi-year frameworks;
- (g) A high percentage of field missions and country analytical work that are carried out jointly with other donors.

4. Effectiveness – Performance-based allocation:

- a) Does the organization use performance-based allocation of resources?

5. Effectiveness – Capacity to manage for development results:

- a) To what extent has the MO taken the results-based management agenda on board? Does the MO have a performance framework based on results and a risk management strategy?
- b) Does the MO have a policy on development effectiveness and does it have a development effectiveness committee?
- c) Does the MO report regularly on corporate performance and are these reports outcomes-focused?
- d) Does the MO report on effectiveness at country levels
- e) Do reports link plans, resource allocations, outputs, outcomes and impacts?
- f) Does Management act on performance results? Are there adequate mechanisms for lesson learning?
- g) Does the MO have an independent evaluation function, i.e. the evaluation function reports to the Board not to Management? How well does the evaluation office function?
- h) Does the institution have a formal and open policy dialogue process with civil society organizations, the private sector and local governments?
- i) Is there an increasing demand from stakeholders on the institution for policy or technical advice?
- j) Is gender equality being implemented? Does the organization have gender-related monitoring capacity? Does the MO disaggregate data by gender? What is the MO's overall reputation in this area?
- k) Is environmental sustainability being implemented as a cross cutting theme? Does the organization have environment-related monitoring capacity? What is the MO's overall reputation in this area?

Annex 9: The Danish PMF

Table 8: Overview of PMF for multilateral development cooperation				
	Focus of PMS	Key Management Documents	Methodology and Source of Information	Reporting and Key Responsible
Corporate Level (Danida)	Measure performance and verify results on overall goals, objectives and indicators in key policy documents	Partnership 2000 MDGs Various sector or issue-specific policies and strategies	Annual report on action plans Assessment of Multilateral Organisations (HQ & Field) Evaluations MOPAN	Annual Performance Report (KVA) Annual report of Danida (UDV)
Organisation level (HQ)	Monitor performance against the objectives, targets and indicators set out in Organisations' own vision and strategy Monitor performance against the objectives, targets and indicators set out in Organisation Strategies	Organisation Strategies & action plans	Multilateral and relevant departments' Assessment of Multilateral Organisations Minutes from High Level Consultations (Dep.) & Annual report on action plans Assessments of RBM Evaluations	Annual Performance Report (KVA) Report to the Board of Danida (Dep.) Annual Performance Report (KVA) Assessment Reports (KVA) Evaluation reports (EVAL)
Country Level (field level)	Monitor performance at country level against the objectives, targets and indicators set out in Organisations' own vision and strategy	Organisation Strategies & action plans	Bilateral Representations's Assessment of Multilateral Organisations MOPAN	Annual Performance Report (KVA) Assessment reports (KVA)
Source: DANIDA (2005b, 17)				

The PFM includes a series of tools that combine qualitative and quantitative information for multilateral and bilateral development cooperation. Key tools for its quality assessment of multilateral organizations include:

Organisation strategies govern the Danish cooperation with the multilateral organisations receiving more than DKK 20 million annually as well as organisations that are strategically important to Danish multilateral cooperation. These strategies outline the overall Danish priorities vis-à-vis the organisation and progress in meeting these objectives is measured annually through *action plans* which contain indicators of performance in four pre-determined areas.

High-level Consultations between the MFA and each of the multilateral organisations take place at least biannually to assess the organisation's progress towards implementing the Danish strategy for the organisation and issues relating to the organisation's handling of

all general and earmarked contributions provided by Denmark. The minutes of these meetings constitute important monitoring information.

Perception analyses of the general performance of multilateral organisations are made by multilateral representations and relevant departments in MFA.

Joint donor assessments of multilateral organisations' effort at country level (MOPAN) are carried out annually in a joint effort between Denmark and like-minded donors.

Perception analyses of the performance of multilateral organisations at country level are carried out by Embassies in programme countries.

Evaluations/assessments of multilateral organisations' performance reporting systems are carried out jointly with other donors to provide the basis for support to strengthening the capacity of the organisations to provide high-quality performance information themselves.

Annex 10: The Dutch MMS

The MMS is a report constructed out of the opinions of embassy staff and their peers. It gives an insight into the general perceptions of a given multilateral that the Dutch government funds. Specifically, embassy staff and peers are asked to comment on the following:

- Degree of embassy involvement or co-operation with in-country multilaterals (either through financing modalities, or intense dialogue);
- Perceived extent to which given multilaterals' mandates and activities are contributing to poverty reduction;
- Whether multilaterals' respective mandates relevant to AEV²³ goals;
- Whether multilaterals' respective mandates relevant to AEV (i.e. education, environment / water, repro health, and HIV/AIDS);
- How satisfactory are respective multilateral organizations' contributions to the improvement of the business climate (i.e. in relation to AEV factors: efforts to strengthen the financial sector, fight corruption, train customs and tax officials);
- How satisfactory are respective multilateral organizations' to the improvement of Good Governance (i.e. improvement of dialogue on human rights, democratic principles, the fight against corruption, the business climate, management of natural resources, illegal migration, and human trafficking, legislative and judicial improvements to government structures);
- How satisfactory are respective multilateral organizations' to the improvement of local governmental responsibilities (ownership), and are the organizations programs aligned with those of the government;
- The total amount of financial aid spend by this organization in the country;
- Size of the embassy's portfolio with respective organizations (e.g. finances, projects);
- How satisfactory are respective multilateral organizations' contributions to the AEV goals;
- How satisfactory are respective multilateral organizations' accountability and financial management;
- How satisfactory are respective multilateral organizations' foci on their own mandate;
- How satisfactory are respective multilateral organizations' co-operation with other donors.

23 AEV is the Dutch development report on co-operation: 'Mutual interests, mutual responsibilities'.

Annex 11: The “New Approach”

Principle	Statement
Evaluation Purpose	— To improve future aid policy, programs and projects through feedback of lessons; and to provide a basis for accountability including the provision of information to the public.
Impartiality and Independence	— The evaluation process should be impartial and independent in its function from the process concerned with the policy making, the delivery, and the management of development assistance.
Credibility	— The credibility of evaluation depends on the expertise and independence of the evaluators and the degree of transparency of the evaluation process. Credibility requires that evaluations should report success as well as failures. Recipient countries should, as a rule, fully participate in evaluation in order to promote credibility and commitment
Usefulness	— To have an impact on decision-making, evaluation findings must be perceived as relevant and useful and be presented in a clear and concise way. They should fully reflect the different interests and needs many parties involved in development cooperation
Participation by donors and recipients	— Consistent with the partnership Principle stressed above, whenever possible, both donors and recipients should be involved in the evaluation process.
Donor Cooperation	— Collaboration between donors is essential in order to learn from each other and to avoid duplication of effort
Evaluation programming	— An overall plan must be developed by the agency for the evaluation of development
Evaluation design and implementation	— Each evaluation must be planned and terms of reference drawn up in order to: define the purpose and scope of the evaluation; describe the methods to be used; identify the standards; determine the resources and time required.
Reporting, dissemination and feedback	— Dissemination and feedback must form a continuous and dynamic part of the evaluation process.

Source: (Danida 2006a, 2)

The New Approach is based on assessment by a panel of professional evaluators. It stipulates that this “peer panel” should comprise representatives from both bilateral and multi-lateral agencies, and beneficiary countries as well as independent evaluation expertise. This is thus a peer review by professionals, and not a peer review between organisations (Danida 2006a, 2).

The process for implementing the “New Approach” is divided into four steps:

- 1) Development and agreement on normative frameworks for assessing independence, credibility and use of evaluation evidence;
- 2) Collection of data, and its analysis, against these normative frameworks;
- 3) Agreement reached on the accuracy of the evidence and findings against the frameworks; and
- 4) Development of conclusions and recommendations

Publikationen des Deutschen Instituts für Entwicklungspolitik

Nomos Verlagsgesellschaft

Messner, Dirk / Imme Scholz (Hrsg.): *Zukunftsfragen der Entwicklungspolitik*, 410 S., Nomos, Baden-Baden 2004, ISBN 3-8329-1005-0

Neubert, Susanne / Waltina Scheumann / Annette van Edig, / Walter Huppert (Hrsg.): *Integriertes Wasserressourcen-Management (IWRM): Ein Konzept in die Praxis überführen*, 314 S., Nomos, Baden-Baden 2004, ISBN 3-8329-1111-1

Brandt, Hartmut / Uwe Otzen: *Armutorientierte landwirtschaftliche und ländliche Entwicklung*, 342 S., Nomos, Baden-Baden 2004, ISBN 3-8329-0555-3

Liebig, Klaus: *Internationale Regulierung geistiger Eigentumsrechte und Wissenserwerb in Entwicklungsländern: Eine ökonomische Analyse*, 233 S., Nomos, Baden-Baden 2007, ISBN 978-3-8329-2379-2 (Entwicklungstheorie und Entwicklungspolitik 1)

Schlumberger, Oliver: *Autoritarismus in der arabischen Welt: Ursachen, Trends und internationale Demokratieförderung*, 225 S., Nomos, Baden-Baden 2008, ISBN 978-3-8329-3114-8 (Entwicklungstheorie und Entwicklungspolitik 2) (im Erscheinen)

Qualmann, Regine: *South Africa's Reintegration into World and Regional Markets: Trade Liberalization and Emerging Patterns of Specialization in the Post-Apartheid Era*, 206 S., Nomos, Baden-Baden 2008, ISBN 978-3-8329-2995-4 (Entwicklungstheorie und Entwicklungspolitik 3) (im Erscheinen)

[zu beziehen über den Verlag oder Buchhandel]

Schriftenreihe bei Routledge

Brandt, Hartmut / Uwe Otzen: *Poverty Orientated Agricultural and Rural Development*, 342 S., Routledge, London 2007, ISBN 978-0-415-36853-7 (Studies in Development and Society 12)

[zu beziehen über den Verlag oder Buchhandel]

Berichte und Gutachten

[Schutzgebühr: 9,63 Euro; zu beziehen direkt beim DIE oder über den Buchhandel. Diese Schriftenreihe des DIE wurde eingestellt und ab November 2004 durch die neue Schriftenreihe „*Studies*“ ersetzt.]

Studies

- 30 *Loewe, Markus et al.*: The Impact of Favouritism on the Business Climate: A Study on *Wasta* in Jordan, 200 S., Bonn 2007, ISBN 978-3-88985-358-5
- 29 *Grävingsholt, Jörn / Claudia Hofmann / Stephan Klingebiel*: Development Cooperation and Non-State Armed Groups, 112 S., Bonn 2007, ISBN 978-3-88985-353-0 (Deutsche Fassung: ISBN 978-3-88985-333-2 – Studie 24)
- 28 *Leiderer, Stefan et al.*: Public Financial Management for PRSP Implementation in Malawi: Formal and informal PFM institutions in a decentralising system, 181 S., Bonn 2007, ISBN 978-3-88985-345-5
- 27 *Altenburg, Tilman et al.*: From Project to Policy Reform: Experiences of German development cooperation, 146 S., Bonn 2007, ISBN 978-3-88985-344-8
- 26 *Chahoud, Tatjana et al.*: Corporate Social and Environmental Responsibility in India – Assessing the UN Global Compact's Role, 118 S., Bonn 2007, ISBN 978-88985-336-3

[Schutzgebühr: 10,00 Euro; zu beziehen direkt beim DIE oder über den Buchhandel.]

Discussion Paper

- 18/2007 *Gu, Jing / John Humphrey / Dirk Messner*: Global Governance and Developing Countries: The Implications of the Rise of China, 25 S., Bonn 2007, ISBN 978-3-88985-356-1
- 17/2007 *Burnell, Peter*: Does International Democracy Promotion work?, 12 S., Bonn 2007, ISBN 978-3-88985-354-7
- 16/2007 *Schirm, Stefan A.*: Die Rolle Brasiliens in der globalen Strukturpolitik, 27 S., Bonn 2007, ISBN 978-3-88985-352-3
- 15/2007 *Dussel Peters, Enrique / Günther Maihold*: Die Rolle Mexikos in der globalen Strukturpolitik, 54 S., Bonn 2007, ISBN 978-3-88985-351-6
- 14/2007 *Müller, Inga*: Perspectives for Germany's Scientific-Technological Cooperation with Sub-Saharan Africa, 54 S., Bonn 2007, ISBN 978-3-88985-350-9 (Deutsche Fassung: ISBN 978-3-88985-328-8 – Discussion Paper 17/2006)
- 13/2007 *Engel, Ulf*: Entwicklungszusammenarbeit mit Nigeria – Eine Analyse aus der Ankerlandperspektive, 60 S., Bonn 2007, ISBN 978-3-88985-349-3
- 12/2007 *Neubert, Susanne et al.*: Poverty Oriented Irrigation Policy in Kenya: Empirical Results and Suggestions for Reform, 54 S., Bonn 2007, ISBN 978-3-88985-348-6
- 11/2007 *Bucar, Maja / Maria João Marques / Anja Mesic / Eva Plibersek*: Towards a Division of Labour in European Development Co-operation: Case studies, 51 S., Bonn 2007, ISBN 978-3-88985-347-9

[Schutzgebühr: 6,00 Euro; zu beziehen direkt beim DIE oder über den Buchhandel.]

Eine vollständige Auflistung erhältlicher Publikationen des DIE finden Sie unter:

<http://www.die-gdi.de>