

Peer Review

UNDP EVALUATION OFFICE

DECEMBER 2005



PEER ASSESSMENT OF EVALUATION
IN MULTILATERAL ORGANISATIONS

United Nations Development Programme

MARY COLE

NIELS DABELSTEIN (CHAIR)

TONY FAINT

TED KLIEST

LUCIANO LAVIZZARI

MINISTRY OF FOREIGN AFFAIRS OF DENMARK
EVALUATION DEPARTMENT

16 DECEMBER 2005

© Ministry of Foreign Affairs of Denmark
January 2006

Production: Evaluation Department, Ministry of Foreign Affairs
Cover illustration: UNDP Evaluation Office
Graphic Production: Designgrafik A/S, Copenhagen

ISBN: 87-7667-387-1
e-ISBN: 87-7667-389-8
ISSN: 1399-4972

This report can be downloaded from www.evaluation.dk

This report can be obtained free of charge by ordering on www.evaluation.dk or from

DBK Logistik Service
Mimersvej 4
DK 4600 Koege
Denmark

Phone: +45 3269 7788

Contact the Evaluation Department: [EVAL@um.dk](mailto: EVAL@um.dk)

Responsibility for the content and presentation of findings and recommendations rests with the authors.

The views and opinions expressed in this report do not necessarily correspond to the views of the Ministry of Foreign Affairs of Denmark, other development partners or UNDP.

Foreword

Who evaluates the evaluator? Occasionally the independence, credibility and utility of evaluation are challenged both by those being evaluated and by the users of evaluation, who may often be the same.

In recent years several major external evaluations of multilateral agencies have been organised, mostly by consortia of funding governments. While valuable for providing evidence of the agencies' contributions to development results, they are indicative of a lack of confidence in the agencies' own evaluation systems to provide credible information. Is that lack of confidence justified? To answer that question the central question to be asked is:

“whether a multilateral agency's evaluation system produces evaluations which are credible, valid and useable for learning and accountability purposes.”

This assessment represents a first test of a new approach, designed under the auspices of the OECD/DAC Network on Development Evaluation.ⁱ It aims at assessing and enhancing *multilateral agencies' own evaluation capacity and performance*, with an ultimate view to improving their development performance. The approach used is a “peer assessment”, and the United Nations Development Programme (UNDP) took the lead by volunteering to undergo the first such assessment.

The Peer Panel comprised:

- | | |
|--------------------------|---|
| <i>Mary Cole</i> | Independent Evaluation Expert. South Africa. Secretary IDEAS. Formerly Head of the Operations Evaluation Unit of the Development Bank of Southern Africa. |
| <i>Niels Dabelstein</i> | Head, Danida Evaluation Department, Ministry of Foreign Affairs, Denmark. Formerly Chair of DAC Working Party on Development Evaluation. |
| <i>Tony Faint</i> | Advisor, DFID Evaluation Department, UK, Formerly Director of DFID International Division. |
| <i>Ted Kliet</i> | Acting Deputy Director, Policy and Operations Evaluation Department, Ministry of Foreign Affairs, The Netherlands. |
| <i>Luciano Lavizzari</i> | Director, Office of Evaluation, International Fund for Agricultural Development (IFAD). |

The team received invaluable assistance from two advisors:
Bernard Wood and *Paul Balogun*.

A Peer Assessment is conducted on a non-adversarial basis, and it relies heavily on mutual trust among the entities involved, as well as their shared confidence in the process. UNDP's Evaluation Office has engaged in this process in an open and constructive dialogue, sharing information, thoughts and ideas with the Panel, so enabling it to come to its conclusions with confidence. Numerous staff of UNDP at all levels, as well as members of the Board, have given their time and insights, greatly facilitating the work of the Panel.

The three crucial aspects of evaluation – *credibility, independence, and utility* – were adopted as the broad yardsticks in assessment, and they are the focus of the main chapters of this report. As the benchmarks for assessing these three broad aspects of the performance of UNDP's evaluation system, a set of thirteen broad Norms for Evaluation in the UN System, formally adopted by the UN Evaluation Group (UNEG) was used. The Panel has made its assessment against these benchmarks as well as against prevailing international practice in bilateral and multilateral development agencies. To inform its judgements, the panel has conducted extensive documentary research, numerous interviews with UNDP staff, Board members and evaluators, both internal and external, as well as intensive discussions with UNDP's Evaluation Office.

The above generic question has been rephrased: "Does UNDP's Central Evaluation Office produce evaluations which are credible, valid and useable for learning and accountability purposes as tested by internationally recognised evaluation peers?"

The short answer to this question is that

"The United Nations Development Programme has an Evaluation Office (EO) that enjoys an acceptable level of independence and which produces evaluations that are credible, valid and useful for learning and strategy formation in the organisation. At the same time, its potential for helping strengthen accountability and performance assessment is being underexploited, both for the purpose of accountability and as an essential basis for learning".

The more detailed answer – as well as suggestions to further enhance the evaluation function and performance assessment in UNDP – can be found on the following pages. While the report is centred on the fairly specialised subject of evaluation, its primary intended audience is one of decisionmakers and other users of evaluation. We hope that UNDP as a whole – its Board of member countries, senior management and staff – will make use of our assessments and suggestions to further strengthen the conduct and use of evaluation in the organisation. We see the current elaboration of a new evaluation policy for UNDP as a unique opportunity to do so. We also hope that this new assessment approach will provide a solid contribution to more effective multilateral evaluation and development performance.

Niels Dabelstein
Chair of the Peer Panel

Abbreviations and Acronyms

ADR	Assessment of Development Results (Country Programme Evaluation)
BoM	Bureau of Management
EB	Executive Board
EO	Evaluation Office of UNDP
GCF	Global Cooperation Framework
IDEAS	International Development Evaluation Association
MDGs	Millennium Development Goals
MYFF	Multi-Year Funding Framework
NEPAD	The New Partnership for Africa's Development
OAPR	Office of Audit and Performance Review
OECD DAC	Development Assistance Committee of the Organisation for Economic Cooperation and Development
OSG	Operations Support Group
PRSP	Poverty Reduction Strategy Paper
RBM	Results Based Management
RCA	Results Competency Assessment
ToR	Terms of Reference
UNDAF	United Nations Development Assistance Framework
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
WTO	World Trade Organisation

Table of Contents

Executive Summary	8
Suggestions to the Executive Board, the Administrator, and the Evaluation Office	12
1. Purpose and Process	15
Ensuring Effectiveness	15
Testing a New Approach	15
Peer Assessment	17
The Current Evaluation System in UNDP: A Sketch	17
2. Independence	19
Roles of the Administrator and Executive Board	19
Separation from Line Management	20
Oversight of Independence and Impartiality	21
Transparent Links between Evaluation Planning and Budget	22
Independence and Impartiality of Evaluators	24
Ensuring Access to Needed Information	26
Freedom of Reporting	26
Tracking Management's Response to an Evaluation	27
Summary of Conclusions and Key Suggestions	28
3. Credibility	29
Evaluation Policy	29
Basic Criteria for Evaluations	30
Competence and Capacity	34
Evaluability	36
Are the EO's Evaluations Impartial?	38
Assuring Quality	40
Summary of Conclusions and Key Suggestions	41
4. Utility	43
The Purposes, Importance and Focus of Evaluation	43
Contributions to Managing for Results	44
The Intention to Use Evaluations	45
Transparency and Consultation	46
Management Response and Follow-up	47
Knowledge, Evidence and Added Value	48
Accessing and Disseminating Lessons	49
The Management of Evaluation	51
Summary of Conclusions and Key Suggestions	52
5. The Summative Judgement	54

Annexes

Annex I: A Description of Evaluation in UNDP	55
Annex II: The Approach and Methodology for this Assessment	62
Annex III: The United Nations Norms for Evaluation	68
Annex IV: A Note on the Draft Evaluation Policy for UNDP	74
Annex V: List of Those Consulted	75
Annex VI: Compilation of Evidence and Findings	78

Endnotes**79**

Executive Summary

Background

- i. This assessment is part of a response to the demand for evidence that development cooperation is making a real difference to development results. While the report is centred on the fairly specialised subject of evaluation, its primary intended audience is one of decisionmakers and other users of evaluation.
- ii. After the experience of several major external evaluations of multilateral agencies – mostly organised in recent years by consortia of funding governments – this assessment represents a first test of a new approach, designed under the auspices of the OECD/DAC Network on Development Evaluation.ⁱⁱ It aims at assessing and enhancing *multilateral agencies' own evaluation capacity and performance*.
- iii. The approach adopted is based on a “peer assessment” mechanism, and the United Nations Development Programme (UNDP) took the lead by volunteering to undergo the first such assessment. The central question to be asked is:

“whether a multilateral agency's own central evaluation office produces evaluations which are credible, valid and useable for learning and accountability purposes as tested by internationally recognised evaluation peers”.

- iv. These three crucial aspects of evaluation – *credibility, independence, and utility* – were adopted as the broad yardsticks in this approach, and they are the focus of the main chapters of this report. As the basis for assessing these three broad aspects, the Panel applied a set of thirteen broad Norms for Evaluation in the UN System, formally adopted by the UN Evaluation Group in April, 2005.
- v. It should be stressed that this assessment is not itself a formal evaluation – it is a less comprehensive and in-depth assessment but, as outlined in Annex II, it adheres to a rigorous methodology applying the key principles of evaluation while taking full advantage of the particular benefits of a peer mechanism. It is explicit that “the final conclusions will clearly be a judgement” by the Panel concerned. As the first exercise of its kind, it has also put particular emphasis on deriving and feeding back experience, lessons learned and possible recommendations for the future.

Overall Assessment

- vi. The United Nations Development Programme has an Evaluation Office (EO) that enjoys an acceptable level of independence and which produces evaluations that are credible, valid and useful for learning and strategy formation in the organisation. At the same time, its potential contribution to strengthening accountability and performance assessment is being underexploited, both in its own right and as an essential basis for learning. (See Box on page 32).
- vii. The Evaluation Office musters the requisite competences at the managerial and professional levels and has clearly been strengthening its role and performance in

recent years. Building on this foundation, and using the new Norms for Evaluation in the UN System, the current testing of a new Policy for Evaluation in UNDP offers an extraordinary opportunity to further strengthen the function and its contribution to UNDP's results.

- viii. In this context we offer a number of interrelated suggestions and options for measures: first, to firmly secure its independence; second, to enhance the accountability and related performance assessment functions; and third, suggestions for more systematic stakeholder participation, management response and follow-up, and proactive, targeted communication of evaluation results.
- ix. While of necessity focusing on the central evaluation function, it has been important to take account of evaluation's place among the various information streams available for the governance and management of UNDP, and how evaluation interacts with the others. In that light, we emphasise the need to reinforce UNDP's results based management systems, and in particular the quality and accessibility of information produced by these systems. We note the current top-level initiative to work out a new accountability framework for the organisation.
- x. At the operational level, we endorse the EO's intention to support an increase in the quality and availability of "decentralised" evaluations (managed by country offices and bureaux) on the actual outcomes of UNDP's projects and other activities. In addition to their intrinsic importance for operational managers, these are the indispensable "building blocks" for a strategic evaluation function to serve properly the needs of the Organisation as a whole.

Independence

- xi. There are now several important openings to formalise the impartial and independent status of the Evaluation Office. While the culture and practice of independent evaluation seems well established, the clarification of reporting lines and some other key arrangements would provide useful protection against any possible infringement, real or perceived, and should be reflected in the new Evaluation Policy.ⁱⁱⁱ First and foremost the Panel suggests a clear and direct reporting line from the Director of Evaluation to the Executive Board to ensure that the Director of Evaluation is directly accountable to the Board rather than to the Administrator.
- xii. While the panel found no evidence that UNDP's systems in this area fall short of protecting the independence of EO staff, the sustained ability of the Director to maintain independence would be strengthened if the Evaluation Policy explicitly stated that the Administrator delegates to the EO's Director authority to make operational decisions concerning the recruitment, promotion and dismissal of EO staff and consultants in accordance with UNDP rules and procedures. Similarly, the Panel considers that independence could be further strengthened if the Director of the EO were appointed by the Executive Board.
- xiii. The forthcoming discussion and approval of the new Evaluation Policy by the Executive Board should be an important step in informing and engaging Board Members on how evaluation may further contribute to their fulfilling their governance responsibilities. For the future, individual evaluation reports and synthesis

reports should be utilised as substantive opportunities for engagement with the Board, bringing together the results and main messages about performance emerging from such work. In particular, we would point to the potential for much greater use of evaluation by the Board to serve as a “reality check” on the organisation’s self-assessments of results.

- xiv. We also suggest that the draft Evaluation Policy be amended to establish clearly that the Director of the EO has the final say on the contents of all reports issued by the Office. This should include the Annual Report on Evaluation and the biennial Report on Development Effectiveness, which should be accompanied by management responses by the Administrator and the tracking of follow-up on evaluation recommendations. The draft Evaluation Policy and reporting arrangements also need to ensure that the Executive Board receives the annual plan of proposed evaluations directly from the Director of Evaluation at the same time as a consolidated budget for this work. In this way the Board can fulfil its key envisaged roles of endorsing UNDP’s evaluation agenda and ensuring that adequate resources are available for conducting evaluation.

Credibility

- xv. Given their current objectives, the credibility of evaluations produced by the Evaluation Office is acceptable and improving in relation to prevailing international practice and the relevant UNEG Norms. In addition to bolstering the accountability function of evaluation (discussed below), credibility could be further strengthened by a number of relatively straightforward measures. The Director, staff and consultants used are professionally competent, and the management of evaluations is basically conducive to quality. While some perceptions of bias within evaluations were identified in the course of our discussions, we conclude that there is not sufficient evidence to suggest that the system as a whole is vulnerable to significant bias.
- xvi. We strongly endorse:
- The draft Evaluation Policy’s intention to mandate a built-in allocation for project and outcome evaluations in operational budgets, which (if conducted in an adequate fashion) will help provide a more solid basis for higher-level evaluations in future;
 - The intention to introduce from early 2006 a set of quality standards and mechanisms to support their use, with a carefully defined role for the EO in both quality assurance and quality control of decentralised evaluations;
 - The decision of the EO to introduce “evaluability assessments” as a normal preparatory step in their evaluations; and
 - The agreement to strengthen and systematise the production, and simultaneous presentation, of management responses for all evaluations presented to the Executive Board.
- xvii. These initiatives hold the potential for addressing some of the weaknesses in the present evaluation approach. We also offer a number of other suggestions aimed at increasing transparency and credibility by:

- Strengthening the elements of performance assessment and accountability – particularly the need to evaluate against intended results – in the draft Evaluation Policy and individual evaluations (while flagging again the problems resulting from the documented weaknesses of UNDP’s systems of managing for results);
 - Clarifying the “rules of the game” for all stakeholders in the draft Evaluation Policy and in guiding documents for each evaluation. For example, the fact that the Evaluation Office takes full responsibility, and often an active editorial role, in shaping the content of evaluation reports should be made unambiguously clear in the new draft Evaluation Policy; and
 - Taking steps to reduce occasional perceptions of a lack of impartiality.
- xviii. Our examination of evaluation practice in UNDP has revealed relatively little emphasis to date on the involvement and “ownership” of partnercountry stakeholders in these evaluations, beyond serving as interviewees, and participants in follow-up workshops where these are held. The UN Norms for Evaluation do not address this as a major issue while the present draft Evaluation Policy for UNDP sets out more ambitious principles, for example through more joint evaluations and capacity building. While the nature of some of UNDP’s thematic evaluations may not be especially conducive to intensive engagement by programme countries, we suggest that the Evaluation Office seeks creative ways of strengthening programme country involvement and partnership in all its work as well as in UNDP’s decentralised evaluation activity. To make this a priority would be consistent with good practice and especially UNDP’s own vocation to support country ownership and capacity building.
- xix. Welcome early steps are now being taken to strengthen UN coordination around evaluation through the UN Evaluation Group. We also assume that the EO will be working to operationalise the other principles featured in the draft Policy around human development and human rights and ethics.

Utility

- xx. The understanding of evaluation and its importance is reasonably strong in UNDP and it has been getting stronger in recent years, including some greater interest at the Executive Board level. Recent UNDP evaluations are geared and timed to decisionmaking and have a growing potential to contribute more during the business cycle, with increased management “take-up”. The evaluations produced are not shelved or treated as irrelevant or unpalatable – they have stirred and informed debate and higher order decisionmaking, even in cases where some of the conclusions remain contested. The trend of use is upward, and new opportunities are clearly emerging, both in thematic and country level evaluations.
- xxi. At the same time, we find that only some of the purposes of evaluation are being properly pursued, and we strongly suggest that the Evaluation Office reexamine and adjust a current imbalance, to give greater weight to accountability as well as learning from evaluation. In particular, after the past three years of experimentation with the current approach to country programme evaluations a definitive model should be adopted as soon as possible, rectifying as far as possible the limitations in the current “goalfree” evaluation approach, and supported by a strong methodology that promotes accountability.

- xxii. Achieving and maintaining the necessary degree of transparency and consultation with all stakeholders at all stages of evaluations is demanding, but without it the base for ownership may not be built, the take-up of results is likely to be weaker, and indeed the due process and good faith in the process can be called into question. We suggest further clarifying mutual expectations among parties on the subject of consultation in the new draft Evaluation Policy as well as drawing up and following a full roadmap for consultation in each evaluation, clearly specifying these arrangements.
- xxiii. To ensure the effective use of evaluation in UNDP – while also reinforcing its credibility and independence – it is hard to overstate the importance of improved arrangements for management response and the systematic follow-up of evaluation recommendations. Building on recent precedents, the draft Policy offers the prospect of a much more thorough system. We suggest that the new practice of responding to recommendations with a management response identifying the unit's responsible and timebound strategies for implementation should be extended to *all* evaluations carried out within UNDP. The relevant managers should be responsible for acting on accepted recommendations, and ensuring that appropriate monitoring and reporting systems are in place to track implementation.
- xxiv. UNDP's current accessibility and disclosure arrangements for the centrally managed evaluations are appropriate. We suggest that "decentralised" evaluations should also be put in the public domain. We further suggest that the Evaluation Office invest energetically in the specialised skills needed to succeed in proactively disseminating findings in userfriendly forms, and we make a number of specific suggestions to this end.
- xxv. Overall, the Annual Report on Evaluation should serve to clarify and reinforce the logic of planning and coverage of the programme of centrally managed evaluations, and explain the process by which the plan is arrived at, as well as highlighting the main results and lessons of centrally and decentrally managed evaluations.

Suggestions to the Executive Board, the Administrator, and the Evaluation Office

To The Executive Board and the Administrator

1. The Panel suggests a clear and direct reporting line from the Director of Evaluation to the Executive Board that would ensure that the Director of Evaluation is directly accountable to and appointed by the Board, rather than the Administrator. While the culture and practice of independent evaluation seems well established, we believe that the clarification of reporting lines and some other issues would provide useful protection against any possible infringement, real or perceived, and should be reflected in the new Evaluation Policy.
2. Given the crucial importance of an informed and engaged Board in securing the independence of evaluation, the discussion and approval of the new Evaluation Policy, may offer important openings for the future. In particular, we would point to the potential of much greater use of evaluation by the Board to provide a "reality check" on the organisation's self assessments of results.

3. The draft Evaluation Policy and reporting arrangements also need to be amended to clarify the linkage between evaluation planning and budgeting, if the Executive Board is actually to fulfil the roles outlined in the draft Evaluation Policy of approving UNDP's evaluation agenda and ensuring that adequate resources are available for conducting evaluation.
4. We also suggest that the draft Evaluation Policy be amended to establish clearly the principle that the Director of the EO always has the final say on the contents of all reports – including the Annual Report on Evaluation – issued by the EO and that there should always be a parallel management response by the Administrator.

To the Administrator

We strongly endorse:

1. The draft Evaluation Policy's intention to mandate a built-in allocation for decentralised project and outcome evaluations in operational budgets, which (if conducted in an adequate fashion) will help provide a more solid basis for higher-level evaluations in future. EO needs to have a carefully defined role in both quality assurance and quality control of decentralised evaluations.
2. The agreement to strengthen and systematise the production, and simultaneous presentation, of management responses and follow-up tracking arrangements for all evaluations presented to the Executive Board. We suggest extending this to all types of evaluations in UNDP.

We further suggest:

3. Strengthening the elements of performance assessment and accountability – particularly the objective of evaluating against intended results – in the draft Evaluation Policy and individual evaluations. This objective is linked to overcoming the problems resulting from the weaknesses of UNDP's systems of managing for results, which will hopefully be improved through the current initiative to work out a new accountability framework for the organisation.
4. Addressing in the new Policy the possibility of evaluation reports specifically reflecting dissenting views where warranted.

To the Evaluation Office

1. The Evaluation Office will need to target carefully the future roles and content of the Annual Report on Evaluation to the Executive Board and the periodic Reports on Development Effectiveness as opportunities for substantive engagement with the Board on the evaluation programme, its overall findings, and tracking the implementation of agreed recommendations.

We strongly endorse:

2. The decision of the EO to introduce “evaluability assessments” as a normal preparatory step in their evaluations.
3. The intention to introduce a set of evaluation quality standards and institutional mechanisms for supporting and utilizing these from early 2006.

We further suggest:

4. Clarifying the “rules of the game” for all stakeholders in the Evaluation Policy and in guiding documents for each evaluation. For example, further clarifying mutual expectations among parties on the subject of consultation and the fact that the Evaluation Office takes full responsibility, and often an active editorial role, in the content of evaluation reports should be made unambiguously clear.
5. Taking steps to reduce occasional perceptions of a lack of impartiality.
6. Seeking creative ways of strengthening programme country involvement and partnership in all its work for example through more joint evaluations and evaluation capacity building. To make this a priority would be consistent with good practice and especially UNDP’s own vocation to support country ownership and capacity building.
7. After the past three years of experimentation with the current approach to country programme evaluations, a definitive model should be adopted as soon as possible, rectifying as far as possible the limitations in the current “goalfree” evaluation approach.
8. Investing energetically in the specialised attention and skills to succeed in proactively disseminating of its findings in userfriendly forms. We make a number of specific suggestions to this end.

1. Purpose and Process

1. This assessment is one part of a response to the insistent demand for evidence that development cooperation is making a real difference to development results. While the report is centred on the fairly specialised subject of evaluation, its primary intended audience is one of decisionmakers and other users of evaluation.

Ensuring Effectiveness

2. In the field of international development assistance, examining and strengthening effectiveness is both more important and challenging than in many other areas. In some ways it is also relatively advanced. As the international community has come together around a shared agenda for improving performance in development and development cooperation, particularly over the past decade, the need for strong monitoring and evaluation has been a key concern. It is closely linked to setting clearer objectives, improving working practices, and applying the lessons of experience – all with developing countries and their people more firmly in the “driver’s seat”, where they must be if durable progress is to be achieved.
3. All the key actors in development cooperation are now part of this drive for evidence of effectiveness and improvement, and they must meet many of the same tests. Alongside bilateral assistance provided directly by governments and non-governmental assistance, the different multilateral development agencies channel a large share of total assistance flows to developing and transitional countries. Their member countries need credible evidence of effectiveness and benchmarks for improvement, as do their partners on the ground, as do their own senior managers and operational staff. At the same time, assessing the effectiveness of these organisations can pose particular challenges because of their wide membership, international governance and sometimes the range and types of their activities.

Testing a New Approach

4. After the experience of several major external evaluations of multilateral agencies – mostly organised in recent years by consortia of funding governments – this assessment represents a first test of a new approach, designed under the auspices of the OECD/DAC Network on Development Evaluation.^{iv} It aims at assessing and enhancing *multilateral agencies’ own evaluation capacity and performance*, and thus helping improve their development performance. It also aims to foster the increased use of the products of the multilateral agencies’ own evaluation functions by stakeholders, including funding governments, and minimise the reliance of funding governments on evaluating the performance of these agencies through external evaluations.
5. The approach adopted is based on assessment against defined and agreed-upon international benchmarks and relevant best practice, and uses a “peer assessment” mechanism to call upon the judgement of a diverse expert panel with a high level

of understanding of these benchmarks and their current application across a range of agencies. The United Nations Development Programme (UNDP) took the lead by volunteering in 2004 to undergo the first such assessment, against the following core question:

“whether a multilateral agency’s own central evaluation office produces evaluations which are credible, valid and useable for learning and accountability purposes as tested by internationally recognised evaluation peers”.

6. These three crucial aspects of evaluation – *credibility, independence, and utility* – were adopted as the broad yardsticks in this approach, and they are the focus of the main chapters of this report.
7. As the basis for assessing these three aspects of the performance of UN organisations, the team was able to apply a set of thirteen broad Norms for Evaluation in the UN System, formally adopted by the UN Evaluation Group in April, 2005, as well as a set of more detailed Standards adopted at the same time. The text of the Norms is found in Annex III. The three key aspects of evaluation have been used to cluster the examination against all these Norms.
8. It should be noted that the new UN Norms, and UNDP’s own draft Evaluation Policy, prominently mention several issues that have not to date received extensive attention under any of these three headings. The draft Evaluation Policy Statement for UNDP sets out several Guiding Principles that “should guide all evaluations in and for UNDP.” Those on National Ownership and Managing for Results are applicable to a wide range of international development agencies. The other two principles that may have special application to UNDP (and presumably other UN agencies) are those around “Human Development and Human Rights” and “UN Coordination and Global Partnership”.^v Attention to these principles has informed the Panel’s assessment, and the findings and implications are discussed in the chapter on Credibility.
9. Once the normative frameworks and process for the assessment were agreed upon with the Evaluation Office (in consultation with the UN Evaluation Group), the succeeding steps were:

Step 2: Collection of data, and its analysis, against these normative frameworks. This took place primarily through extensive documentary research, with a great deal of help from the Evaluation Office, followed by structured and semi-structured interviews with a range of participants and intended users of evaluations. This work was concentrated around five recent evaluations of different types;

Step 3: Agreement by the Panel and EO on the accuracy of the evidence and findings against the frameworks, and on the basis for arriving at judgements;

Step 4: Development of conclusions and recommendations by the Panel, dialogue on the draft report with the Evaluation Office in a mutual learning process, and finalization of the Panel’s report.

Peer Assessment

“Peer review can be described as the systematic examination and assessment of the performance of an entity by counterpart entities, with the ultimate goal of helping the reviewed entity improve its policy making, adopt best practices, and comply with established standards and principles. The examination is conducted on a non-adversarial basis, and it relies heavily on mutual trust among the entities involved in the review, as well as their shared confidence in the process.”^{vi}

10. The device of professional peer assessment, as illustrated in international processes such as those of the OECD, WTO and more recently NEPAD, is already well tested in the development field and brings a number of special strengths to this undertaking. Focusing on the core assessment question, the Panel and its advisors assembled and organised the fullest information base practicable, brought together and tested through direct interviews by the Peer Panellists.^{vii} This base of evidence and findings then provided a basis for drawing conclusions; reflecting the diverse professional experience and judgements of the Peer Panellists themselves, deliberating as a group on an extensive set of subquestions. Because of the Panel members’ comparative experience, a major strength of the peer approach is therefore that their judgements were grounded in a sense of realism. Although the Panel worked to arrive at consensus judgements where possible, there were instances where their judgements were different and a variety of possible suggestions were explored. More generally, while the Panel drew on the experience and knowledge of its members to outline some suggestions, pointers and options for action, it was not the Panel’s role to prescribe any particular model for evaluation in UNDP, thus recognizing that each organisation must ultimately decide for itself which approach is best suited to its particular circumstances.
11. It should be stressed that this overall exercise was not itself a formal evaluation.
 - It was a less comprehensive and in-depth assessment but, as outlined in Annex II, it adhered to a rigorous methodology applying the key principles of evaluation while taking full advantage of the particular benefits of a peer mechanism. It is also explicit that “the final conclusions are a judgement” by the Panel concerned. As the first exercise of its kind, the assessment placed particular emphasis on deriving and feeding back experience, lessons learned and possible recommendations for the future.

The Current Evaluation System in UNDP: A Sketch^{viii}

12. Evaluation is a longstanding and widespread function in UNDP, with a total of 200+ evaluation studies being carried out in 2004. The vast majority of this work was commissioned on a “decentralised” basis by country offices and other operational units and carried out by consultants. As the custodian of the evaluation function in the organisation, UNDP has an Evaluation Office (EO) which is an independent unit reporting to the Executive Board through the Administrator. It should be stressed that this assessment is focused on the work of the central evaluation office – including its relationships with the many evaluations carried out elsewhere in the organisation. However it does not attempt any direct appraisal of the far-flung and diverse activity carried out in UNDP’s “decentralised” evaluations.

13. The Office has responsibility for undertaking and reporting on strategic evaluations (seven in 2004); promoting the use of evaluation findings, lessons and recommendations; ensuring that evaluation findings are accessible; developing evaluation guidelines, methods and methodological tools and systems; assisting in the development of evaluation capacities; generally overseeing the evaluation function in UNDP; and preparing the Annual Report of the Administrator on Evaluation to the Executive Board. It also participates in the professional networks on evaluation within the United Nations system and more broadly. The Office currently has 11 full-time professional evaluation staffmembers, in addition to the Director and Deputy Director, and a total annual expenditure of USD 5.4 million.

14. As of September 2005, UNDP has a draft "Evaluation Policy Statement" for the organisation which the Administrator has endorsed, with a view to final approval by the Executive Board in 2006. This statement reflects the new UN Norms, and for the first time pulls together in one instrument all the guiding principles and norms, key evaluation concepts, main organisational roles and responsibilities involved. It defines the types of evaluation covered, identifies the key elements of a system for learning and knowledge management, and outlines the capacity and resource requirements to enhance excellence in the development of a culture of evaluation and learning throughout the organisation, and the need to deepen partnerships in evaluation. Because the draft policy statement is very new and not finally adopted, it cannot yet be taken by this assessment as an established reflection of the current performance of the UNDP evaluation system. At the same time, the Panel's conclusions and suggestions recognise it as a very important indication of the organisation's future directions, as well as a reflection of a good deal of UNDP's existing policy and practice.

2. 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.”^{ix}

Roles of the Administrator and Executive Board

Do the Executive Board and the Administrator foster an enabling environment for evaluation? (Norms 2.1 & 2.2)

15. The EO’s current mandate states that:

“The mandated responsibility of the Evaluation Office is to support the Administrator in his substantive accountability function and contribute to organisational learning through provision of systematic and independent assessment of results, effectiveness and impact of the substantive activities of the programme, including the special purpose funds under the Administrator’s responsibility.”

16. This provides a clear statement that the key responsibility in fostering the EO’s role lies with the Administrator. To date there has been no proactive role of the Executive Board in fostering an independent EO or oversight of the evaluation function in any systemic way. In practice, however, the Executive Board affects the environment for independent evaluation through two other routes:

- First, by requesting specific evaluations and progress reports. The evidence shows that the Executive Board is directly requesting an increasing number of evaluations and progress reports. While the reasons why the Board requests specific evaluations vary, this does potentially act to strengthen the EO’s ability to evaluate issues that may be sensitive to senior management within UNDP.
- Second, by approving specific suggestions that are raised in the ‘Annual Report of the Administrator on Evaluation’. Recent examples of this were approval of the codification of an evaluation policy for UNDP during the 2005 reporting period and the formalisation of a system for tracking evaluation recommendations.

17. Many of those interviewed during this assessment suggest that while there has not been a formal policy statement in place, there has been a *de facto* evaluation policy in operation. This is based on custom and emerges from a set of organisational norms and practices that are reinforced by the maintenance of a professional cadre of evaluation staff within the EO. For example, recent Administrators have been supportive when the EO proposes actions to bolster its independence.

18. A draft version of the Evaluation Policy was endorsed by the UNDP Senior Management Team in early September 2005 for roll out and field testing of new approaches. This will then provide the basis for refinement and presentation for endorsement to the Executive Board at their next annual session in June 2006, as the Board has requested.

Conclusions

19. We believe that the present draft Evaluation Policy does clarify the roles and responsibilities of the Administrator and Executive Board in terms of their respective roles and responsibilities. It coincides with a rising level of interest in the potential contribution of this function to the effective management and governance of UNDP.

Suggestions

20. Particularly with the Executive Board, the discussion and approval of the new Evaluation Policy, following on from the review and discussion of a number of evaluations by the Board in the recent past, may open up important new opportunities for informing and engaging the Board more actively around evaluation. For the future, the EO needs to consider how best to communicate with the Board on three main sets of issues, taking account of the need for brevity in documents presented to the Board. These are (i) the EO's work in the past year and its plan of evaluations and budget for the coming year; (ii) a report that consolidates results and impact and main crosscutting issues relating to UNDP's performance as evidenced by evaluations; and (iii) a report that summarises whether management are implementing agreed evaluation recommendations on schedule. Doing this will require the EO to target carefully the future roles of the Annual Report on Evaluation to the Executive Board and the periodic Reports on Development Effectiveness as opportunities for substantive engagement with the Board.

Separation from Line Management

Is the EO located independently from the other management functions so that it is free from undue influence? (Norm 6.1).

21. Yes. The EO is directly accountable to, and reports on, evaluation results to the Administrator and therefore supports the Administrator in his substantive accountability function. Functionally, the EO is a separate department within UNDP which reports for administrative purposes to the Associate Administrator. As such, it is located organisationally outside the staff or line management function of the programmes, activities or entities being evaluated. The EO has also restricted itself to an observer role in senior decisionmaking committees, as a further measure to ensure that it is free from potential conflicts of interest in evaluating senior management decisions.
22. The model proposed in the draft Evaluation Policy specifies that the Administrator has the primary responsibility for ensuring EO independence, while the EO is accountable to, and reports on, evaluation results to the Administrator and through the Administrator to the Executive Board. This reflects present practice in UNDP.

Conclusion

23. Our major conclusion is that there is no evidence that the independence of the EO has in fact been compromised to date by its location or the fact that it reports to the Administrator. At the same time, however, avoiding such risks under the current arrangement will always be dependent on the understanding and forbearance

of the Administrator of the day, and the perceived risk of self-censorship will always be present. Given the character of this function, we take the view that its independence – real and perceived – would be best secured by the Director of Evaluation being directly accountable to the Executive Board rather than the Administrator.

Suggestions

24. We suggest that the perception, and sustainability, of independence be enhanced by providing in the draft Evaluation Policy for a clear and direct reporting line from the EO to the Executive Board. In practical terms this would mean that the Evaluation Policy would state that:
 - The Director of the EO would have the authority to present evaluations and the Annual Report on Evaluation and the periodic Report on Development Effectiveness to the Executive Board without prior clearance of the contents from the Administrator or any other official within UNDP;
 - The evaluation programme, individual evaluations and the Annual Report on Evaluation would be presented to the Executive Board, and subsequently published, in the name of the EO;
 - A management response, in the name of the Administrator, would be presented to the Executive Board at the same time as evaluations and the Annual Report on Evaluation.
25. This approach would also assist the Administrator, by removing a potential conflict of interest between ensuring the independence of the EO and also ensuring that there is an adequate management response to evaluation results.
26. However, there is a view within the Panel that the UNEG Norms themselves are not sufficiently clear on this principle, and the Panel points to the examples of the World Bank, Inter-American Development Bank, the International Fund for Agriculture Development and other organisations, where the Evaluation Director is appointed by, and fully responsible to the Executive Board, rather than management. Relatedly, the practice within several international financial institutions of requiring that the Director of Evaluation not move subsequently into an operations position in the organisation might also be considered as a support to maintaining the independence of future Directors of the EO. However the trade-offs of restricting in-house selection to end-of-career candidates must also be taken into account when considering such a practice.

Oversight of Independence and Impartiality

Do the Executive Board and Administrator ensure that evaluations are conducted in an impartial and independent fashion? Do they ensure that evaluators have the freedom to conduct their work without repercussions for career development? (Norm 2.4 & 6.2-6.5).

27. The Panel's discussions with Executive Board members suggest that there may not yet be a broad consensus within that body on how independent the evaluation function should be and how deeply the Board should be engaged. Unlike in a number of other multilateral organisations, for example, UNDP's Executive Board does not have a subcommittee with special responsibility for evaluation. At this

point, evaluations mandated by the Board are presented directly to the Board by the Director of Evaluation, while others may be discussed informally, but the Executive Board has limited opportunity and means to ensure that evaluations are carried out in an impartial and independent fashion.

28. Internally, direct reporting to the Administrator and insulation from decisionmaking committees are the main measures that specifically aim to reinforce the independence of the EO. Technically, there should be no fear by evaluation managers of adverse repercussions on career prospects, since promotion or reassignment is based upon personal performance under the corporate Results Competency Assessment (RCA) and the Director is responsible for the assessment under the RCA of all EO professional staff. We have found no evidence of any constraints in this area, and in fact observed the recent promotion outside the Evaluation Office of the manager of an evaluation highly critical of a key senior management strategy. Equally, even those consultants who derive a large share of their work from UNDP observe that they have not been pressured to tailor their conclusions. However, we note that the Director's own RCA is currently carried out by the Administrator, Associate Administrator and the Senior Management Team, which holds potential for adversely affecting the Director's subsequent career path within the organisation.

Conclusion and Suggestions

29. While the culture and practice of independent evaluation seems well established, we believe that the clarification of reporting lines as above and some other suggestions below would provide useful protection against any possible infringement, real or perceived, as well as strengthening the value of this instrument to the Executive Board. We believe that these kinds of measures should be explicitly reflected in the new Evaluation Policy.

Transparent Links between Evaluation Planning and Budget

Are the EO budget and plan of evaluations linked so that it is clear that adequate resources are allocated to enabling the EO to operate effectively and with due independence? (Norms 2.3, 2.6 & 4.2).

30. Following a canvassing and consultation process that is similar to that found in many other organisations, the Director of the EO has the discretion to select which evaluations will be programmed. These are then identified in a biennial rolling plan, which is posted on the EO's website. The biennial plan of evaluations is received, but not explicitly endorsed, by the Executive Board. Evaluations which will be completed or initiated in the coming year are identified in the 'Annual Report of the Administrator on Evaluation' that is presented to the Board each June.
31. Thematic evaluations led by the EO are not, nor have ever been, intended to have full coverage of all UNDP supported activities. The evaluation function, unlike in some organisations, is also not mandated to validate directly the self-assessments of performance against the strategic results framework (the Multi-Year Funding Framework) used in UNDP's RBM system. Rather it is intended as a strategic instrument evaluating key corporate, thematic and country issues as determined by the Executive Board, senior management, country counterparts and the EO.

In terms of thematic evaluations, the evidence suggests that the EO has had the right to select which topics will be evaluated, without being compromised by any pressure.

32. The situation for the country programme evaluations is slightly different, with the selection of country programmes being partially driven by the active interest of the bureaux and/or individual country offices. The EO's position is that it has not required the final say in selecting which specific country programmes are evaluated. During the past three years the primary focus of the country programme evaluations has been on supporting learning and testing the evaluation methodology, and selection may become more of an issue as the ADR (Assessment of Development Results) approach is finalized and the evaluation programme expanded.
33. The EO's current total budget is a combination of funds from five separate sources. Within this, the EO has found that the core budget has provided a sufficiently predictable base over the past five years for the existing programme of evaluations.
34. The basic EO budgeting process follows the same rules and procedures as used for other parts of UNDP. The Bureau of Management (BoM) proposes the programme budget on a two-year basis, which after discussion is agreed by the Executive Board at the September Board meeting. In order to increase its budget, the EO must present a case for an increase in programme funding which will be assessed, along with all other requests for additional programme funding, by BoM. When assessing the EO's bid for additional programme funding, BoM looks at past funding levels, the number of staff, and the number of evaluations that have been programmed.
35. The Executive Board approves the most significant components of the EO's budget on a biennial basis, as part of its approval of the overall UNDP budget. However, the separate components of the EO budget approved are not consolidated in any single place within the overall budget. The EO budget is not discussed in the context of the EO's proposed plan of evaluations by the Executive Board.
36. While the EO suggests that the overall budget available is adequate to support the current evaluation portfolio, the level of financial resources available may be a constraint for some individual evaluations and the possibilities for expanded coverage. Precise comparisons of the budget for individual EO evaluations with those from many other agencies are difficult because of differences in the scope and character of these evaluations and the extent to which evaluation office staff participate directly in the evaluations. However, in the case of the Assessments of Development Results (country programme evaluations) the available evidence suggests that the average UNDP evaluation budget lies towards the lower end of the range across development agencies.^x It should also be noted that under the present proposal for expanding evaluation to cover all country programmes, two forms of country programme evaluation are anticipated. Annually, for between eight and ten country programmes, a process similar to the present ADR approach would be used. For the remaining country programmes reaching the end of their programme cycle, around 15 programmes, a light-touch evaluation would be carried out. We note that the envisaged average budget for these new "light-touch" evaluations would be far below that of comparable evaluations elsewhere, and would be comparable to that of an outcome evaluation commissioned by a UNDP Country Office. This proposal to expand coverage can also be contrasted with practice in most other agen-

cies, which do not attempt to carry out 100 per cent evaluation coverage of their country programmes at the end of cycle.

Conclusions

37. The Panel's overall conclusion is that the links between the planning and budgeting of the evaluation programme are not very clear and that the Executive Board has neither the opportunity, nor the required information, to take an informed position on whether or not the EO's budget is adequate. This lack of direct linkage between the programme and budget does not yet appear to have unduly steered the work programme. We also conclude that to carry out the expanded programme of evaluations and quality assurance and control functions identified in the new Policy will require a significant increase over the present budget envelop and raises the danger of whether funding will be sufficient – particularly for country programme evaluation – to maintain the credibility of evaluations carried out.

Suggestions

38. The draft Evaluation Policy gives the Executive Board the role of approving UNDP's evaluation agenda and ensuring that adequate resources are available for conducting evaluation. If the Board is to discharge this function, it should receive a full picture from the Director of Evaluation of the EO's budget alongside the proposed work programme.
39. The Evaluation Office is not exempt from some of the same uncertainty and need for voluntary funding demanded of their colleagues throughout the organisation, and this makes it difficult to present the firm, consolidated budget that would be the ideal. However, it is suggested that the EO, at a minimum, should present a consolidated budget when seeking approval of its Evaluation Agenda by the Executive Board. Discussion of the budget at this point would then strengthen its ability to ensure appropriate funding during negotiations with UNDP's Bureau of Management.

Independence and Impartiality of Evaluators

*Do evaluators operate in an independent and impartial manner? (Norm 2.5, 5.3, 6.3 & 6.4)
See also Paragraphs 98-108 below.*

40. One possible source of bias would be if evaluation managers or consultant evaluators were involved in evaluations of programmes which they had either planned or managed or where they might do so in the near future. The Panel has encountered no evidence of evaluators evaluating programmes which they had previously planned or managed. However, there is evidence that avoiding the possibility of working with programmes that have recently been evaluated by the EO might be seen as a concern for both evaluation managers and contracted evaluators.
41. Such situations may become increasingly common if the number of country programmes evaluated increases, as is planned, since EO senior evaluation managers can move on to becoming Resident Representatives (two cases in 2005). This may also become an increasing concern for consultant evaluators as illustrated in the

recent case of an evaluator being asked by the operational Bureau to take an active role in developing the new strategy whilst also evaluating the current strategy. Particularly given the premium currently placed by the EO on contracting evaluators who have a very detailed knowledge of UNDP, the potential for conflicts of interest will recur.

42. The Director has substantial responsibility for selection and performance appraisal of professional staff within the EO, which is vital given that independence depends on the Office maintaining an internal cadre of competent evaluators. Recruitment and search processes vary somewhat according to the source of funding for the post, and the nature of contract issued. However, the Director and Deputy Director of the Office have the final say in setting the technical requirements for posts and in short-listing candidates. Where it might be relevant, the EO reports that there have been no instances in recent years of being compelled to accept unsuitable staff from elsewhere within UNDP, while the majority of current professional staff have been recruited from outside of UNDP, based on an assessment of their evaluation skills.
43. All staff of the EO are UNDP staffmembers, and are thus subject to the normal Results Competency Assessment procedures for assessing staff performance and internal deployment. In principle, this should mean that working in the Evaluation Office should not adversely affect future opportunities on compensation, training, tenure and advancement within UNDP, and we have not heard of any reported problems on this front. Staff within the Evaluation Office are also formally shielded from external political pressures, so allowing them to report freely. UN staff rules prohibit staff from accepting remuneration, benefits, favours, or gifts of significant value from any government, or other external organisation or person.

Conclusion and Suggestions

44. While the Panel found no evidence that UNDP's systems in this area fall short of protecting the independence of EO staff, the sustained ability of the Director to maintain independence would be strengthened if the Evaluation Policy explicitly stated that the Administrator delegates to the EO's Director authority to make operational decisions concerning the recruitment, promotion and dismissal of EO staff in accordance with UNDP rules and procedures.
45. The practice of excluding evaluation managers or consultant evaluators from evaluations of programmes which they have either planned or managed in the past is essential. With respect to the possibility of evaluators, and especially consultants, going on to some future involvement in programs that they have helped evaluate, we are inclined to be less categorical. There may be occasions where the Director of Evaluation might judge such a role for a consultant to be highly beneficial and quite proper at the appropriate time.
46. With respect to staff members of the Evaluation Office, we find it important that there be at least a sufficient "cooling-off" period before moving to work in an area that one has evaluated.

Ensuring Access to Needed Information

Does the independence of EO impinge on the access that evaluators have to information on the subject being evaluated? (Norm 6.5).

47. There is no evidence that the perceived independence of the EO has led to restricted access to information. Some restriction of access was observed in a very limited number of cases, but this was by all accounts related to concerns of individual managers whose programmes were being evaluated rather than being a symptom of any widespread restrictions on access.

Freedom of Reporting

Does the EO have full discretion in submitting directly its reports for consideration at the appropriate level of decisionmaking pertaining to the subject of evaluation? (Norm 6.1).

48. The degree to which the EO has discretion directly to submit its reports for consideration at the appropriate level of decisionmaking varies. Thematic evaluations that are requested by the Executive Board are invariably submitted to the Board. For thematic evaluations and the country programme evaluations that are not directly requested by the Board, the EO can hold informal consultations with the Board and ensure that copies of the evaluation report are available at Board meetings, as has been done with country programme evaluations.
49. There appear to be no barriers to the EO presenting evaluative evidence to appropriate audiences within UNDP itself. The present draft Evaluation Policy states that:

“The Director of EO is responsible for reporting on evaluations to the EB and for authorizing the dissemination of EO’s evaluation reports and related material in accordance with the disclosure policy of UNDP.”

50. UNDP’s Disclosure Policy states that all UNDP corporate evaluations published by the Evaluation Office are made public through the EO website.

Conclusion

51. While we have not seen evidence of constraints on freedom of reporting in practice, it is a potential concern that neither the disclosure policy nor the draft Evaluation Policy directly specifies the principle that evaluations produced by the EO will not be subsequently amended. The current arrangement for disclosure of centrally managed evaluations is satisfactory, while the status of those carried out under the authority of bureaux or country offices is less clear.

Suggestions

52. We suggest that the draft Evaluation Policy be amended to establish clearly the reporting principle that the Director of the EO always has the final say on the contents of reports issued by the EO. We specifically suggest that – in addition to the direct submission of the Evaluation Work Programme, Annual Report and Report on Development Effectiveness to the Executive Board – all reports of country pro-

gramme evaluations should be provided to the Board for information, and should accompany new country programme strategies when they are presented. We further suggest that all evaluation reports, centrally managed or decentralised, should be placed in the public domain (e.g. on the relevant website) accompanied by management responses, to promote both openness and quality. It may, however, be necessary to restrict publication on the Evaluation Office's own website or use by the Evaluation Resource Centre or the Evaluation Network to those products meeting the quality standards that are to be established in early 2006.

Tracking Management's Response to an Evaluation

Does the EO have the independence to track follow-up of management's response to an evaluation? (Norm 6.2)

53. The Director of Evaluation has the freedom to track whether or not recommendations are implemented, based upon an Executive Board decision from June 2005 which approved the formalisation of a system for tracking evaluation recommendations. The EO is currently considering how to implement this decision of the Executive Board. Two issues when putting such a system in place are:
- Management responses were produced for both the PRSP and MDG evaluations in 2004, but these were not detailed enough to allow tracking of implementation. They did not specifically identify who would be responsible for implementation, with time-bound benchmarks, that could then be used to ensure that implementation was on track. The EO has now developed a standard management response format, that is to be used for future evaluations and identifies for each key recommendation:
 - Management Response to recommendation;
 - Strategy for addressing issue, including indicators of progress, partnerships and timeframe.
 - Second, the EO acknowledges that it does not have sufficient staff to check proactively on the degree of implementation of recommendations and would therefore either have to rely upon operational units to provide the needed data (which has not been agreed) or rely on internal management information systems, which are not necessarily designed to track the relevant actions.

Conclusion

54. While the Evaluation Office has been free to track implementation, neither the necessary procedures nor resources have been in place to do so. The new draft Evaluation Policy will do much to strengthen this system. We suggest that the requirement for management to produce and track implementation of responses, and the ability of Evaluation Office to consolidate and report on the results, should ultimately be extended to all evaluations undertaken in UNDP.

Summary of Conclusions and Key Suggestions

55. The Panel's overall conclusion on issues of independence is that there is no evidence since 2002 that either the Executive Board or UNDP's senior management have attempted to curtail the necessary independence of the EO. The Peer Panel also believes that the present systems, approaches and behaviours in most cases acceptably meet the relevant UNEG Norms for independence, although it has identified a number of steps that would formalise the impartial and independent status of the Evaluation Office and help dispel any possible perceptions that this might be deficient.
56. Given the crucial importance of an informed and engaged Board in securing the independence of evaluation, the discussion and approval of the new Evaluation Policy, following on the scrutiny of a number of evaluations by the Board in the recent past, may offer important openings for the future. The Evaluation Office will need to target carefully the future roles and content of the Annual Report on Evaluation to the Executive Board and the periodic Reports on Development Effectiveness as opportunities for substantive engagement with the Board on the evaluation programme, its overall findings, and tracking the implementation of agreed recommendations.
57. The Panel suggests a clear and direct reporting and accountability line from the Director of Evaluation to the Executive Board. While the culture and practice of independent evaluation seems well established, we believe that the clarification of reporting lines and some other issues, including the appointment of the EO's Director as well as the latter's authority concerning the management of the EO's staff, would provide useful protection against any possible infringement, real or perceived, and should be reflected in the new Evaluation Policy.
58. The draft Evaluation Policy and reporting arrangements also need to be amended to clarify the linkage between evaluation planning and budgeting, if the Executive Board is to actually fulfil the roles outlined in the draft Evaluation Policy of approving UNDP's evaluation agenda and ensuring that adequate resources are available for conducting evaluation.
59. We also suggest that the draft Evaluation Policy be amended to establish clearly the reporting principle that the Director of the EO always has the final say on the contents of reports issued by the EO. This should include the Annual Report on Evaluation, with parallel management responses by the Administrator.

3. 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 successes as well as failures. Recipient countries should, as a rule, fully participate in evaluation in order to promote credibility and commitment.”^{xi}

Evaluation Policy

Does UNDP have an evaluation policy that explains the concept of evaluation, roles and responsibilities and how evaluation evidence will be used? (Norm 3.1).

60. Prior to formal adoption in the coming months of the Evaluation Policy now being tested in draft form, only some of the issues that would be included in an Evaluation Policy for a central evaluation function have been formally agreed and mandated within UNDP. On the other hand, for evaluations managed by operational bureaux and country offices, roles and responsibilities for monitoring and evaluation, evaluation concepts and how to use evaluation evidence have been quite thoroughly documented since 2002 in UNDP’s “Handbook on Monitoring and Evaluating for Results”.
61. The sections of the draft Evaluation Policy, as currently proposed, touch on all of the issues specified in the UNEG Norms, as shown below:
 - a. Explanation of the concept and role of evaluation within the organisation;
 - b. Explanation of the various types of evaluation applied within the organisation (self-evaluation; independent evaluation; centralised/decentralised forms of evaluation; etc);
 - c. Differentiation of evaluation from other types of assessment carried out within the organisation;
 - d. Definition of the roles and responsibilities of the evaluation professionals, senior management and programme managers;
 - e. Emphasis on the need for adherence to the organisation’s evaluation guidelines;
 - f. Explanation of how evaluations are prioritized and planned;
 - g. Description of how evaluations are organised, managed and budgeted;
 - h. Emphasis on the management response for the follow-up of evaluations.

Conclusion

62. The degree to which the proposed Evaluation Policy effectively deals with these issues is discussed in more detail throughout this report and therefore is not repeated here. In general, our conclusion is that the Policy, as currently drafted, would substantially strengthen the base for an effective evaluation function in UNDP, but that the opportunity should not be missed to reinforce and clarify it in several important areas identified under the relevant headings in this text.

Basic Criteria for Evaluations

Do the EO's evaluations meet the criteria identified in the UNEG definition of an evaluation? (Norms 1.2 & 1.4).

63. The UNEG Norm N1.2 specifies that an evaluation:

'...focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof. It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions and contributions of the organisations of the UN system.'

64. Interviews with UNDP senior managers confirmed that in principle they were quite clear on the differences between evaluation and other forms of assessment. An EO assessment of monitoring and evaluation at the country programme level in 2005 also found that there is a common understanding of these differences at programme level.^{xii} By contrast, at the top governance level of the organisation, discussion with Executive Board members from member countries showed that they were not necessarily clear on the differences between the different forms of assessment and how they should each be used.

65. The evidence suggests that EO evaluations produced between 2002 and 2005 meet most, but not all, of the criteria for an evaluation identified under UNEG Norm 1.2 above. While it may not be essential for every evaluation to include all of these elements, it would normally be expected that evaluation reports, at least in their methodology section, would identify and justify which of them will and will not be covered.

66. Review of the Terms of Reference for all evaluations that will have been completed between 2002 and January 2005 showed that all but one,^{xiii} were initially designed to identify the results (accomplishments) of UNDP. The examination of the ToRs for all of the country programme evaluations and the six thematic evaluations^{xiv} that will have been presented to the Executive Board between 2002 and January 2006 shows that:

- EO managed evaluations do set out to examine relevance and effectiveness;
- An examination of sustainability was included in the ToRs for all country programme evaluations but was not included in the ToRs for any of the six thematic evaluations;
- No EO managed evaluation, even those focused explicitly on UNDP programmatic activities, set out to examine efficiency;
- The Thematic studies (four) completed in 2005/06 and all country programme evaluations do attempt to identify development results and UNDP contributions to these results (impacts).

67. In terms of assessing actual results achieved, both the evaluation managers and evaluators involved in the reference cases stated that the benchmarks and criteria used to evaluate performance were in practice mostly developed by the evaluation consultants. However, the evaluation reports did not always clearly state what these benchmarks and criteria were, and there is little indication that evaluators ensured that these benchmarks and criteria were agreed with those responsible for operations being evaluated. This finding was confirmed by most of the operational staff interviewed during this peer assessment.

68. The picture is also mixed about whether or not evaluations have met the criterion of examining whether the *intended* UNDP accomplishments (results) were achieved or not. This is basically an assessment how well UNDP has succeeded in meeting its operational objectives. In terms of the six thematic evaluations that will have been presented to the Executive Board between 2002 and January 2006, in two cases the ToRs included examination of whether the intended results were accomplished. In four cases, the ToRs did not include a requirement to evaluate whether or not the intended results were achieved. In the cases of the current, experimental approach to country programme evaluations (the "Assessments of Development Results") the EO's working methodology is unclear and inconsistent on the extent to which they should assess whether intended results have been achieved. The commendable aim in this new approach was to shift to assessing UNDP's relatively modest activities in the overall context of the country's development objectives and results, rather than evaluating them in isolation, and to take into account unintended results as well as UNDP's chosen objectives. This shift is obviously in line with good practice internationally, but presents all cooperation agencies with the challenge of working out the right level and approach for assessing rigorously their own contributions within the broader context. The EO will be revising its guidance on country programme evaluations in early 2006, but the intended emphasis and approach for examining whether or not intended achievements have been met are not yet clear.

Conclusions

69. While there may be reason for EO managed evaluations not to focus major attention on evaluating efficiency and sustainability, a proper evaluation methodology should explain how these issues are being treated and why. Credibility would also be enhanced if the EO's evaluation reports always provided an explanation of the benchmarks that were used by evaluators and evaluation managers when assessing the actual performance of UNDP.
70. As specified at the beginning of Norm 1.2, a cornerstone for any evaluation office that fulfils its accountability and oversight function is to identify explicitly the differences between intended and actual results achieved as found in the organisation's results reporting and its evaluations. Our discussion with Executive Board members and our experience with evaluation in other organisations shows that information on whether or not the organisation did what it set out to do, and the degree to which self assessments of such performance are reliable and robust, is a key product from a central evaluation office for both an organisation's own senior managers and those to whom the organisation is accountable.
71. A fundamental concern is the results framework and information base that underpin EO's higher-level evaluations. This concern is hardly new to either the organisation or the Executive Board, as can be seen in the persistent requests by the Executive Board in the area of performance reporting by UNDP. Weaknesses in this area are symptomatic of a wider problem in the organisation that has led to the recent top-level initiative to work out a new accountability framework for UNDP as a whole. For the purposes of evaluation, at present there are deficiencies both in the organisation's central monitoring and reporting system and the decentralised evaluation system, and it is not feasible to cost-effectively compensate for these deficiencies within the scope of a higher-level evaluation (e.g. by using or commissioning evaluability and local consultancy studies). The EO's evaluations, while

Evaluation for Accountability and Learning

The UN Norms for Evaluation set out two main purposes for an evaluation:

Accountability: *“It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof. It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions and contributions of the organisations of the UN system.”*

Learning: *“An evaluation should provide evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes of the organisations of the UN system and its members.” (Norm 1.2)*

In the light of its findings about how these two purposes have been treated to date in UNDP, the Panel concluded that there would be value in specifically clarifying international good practice in how these two purposes should be pursued and linked. We have found that the understanding and application of the accountability function of evaluation in UNDP has been weaker than it should be. Other systems are allocated greater responsibility for assuring accountability, while the very fact of producing and conveying an evaluation to the Governing Board and/or senior management has been treated as serving the accountability function.

Our findings from all sources have confirmed, however, that much greater emphasis has been placed by all concerned in UNDP on trying to aim directly for “learning” benefits from evaluation, without giving sufficient attention to accountability as a necessary first step in the learning process. In the evaluation of development cooperation, accountability refers to the systematic assessment of both *expected and achieved* development results, the impact of development assistance and the performance of the parties involved.^{*} In order to realize more fully the combined benefits from these two interrelated purposes, we encourage UNDP’s Evaluation Office to focus on strengthening this substantive accountability purpose in its methodologies, and in the design, conduct and reporting of its evaluations.

We would encourage the vigorous implementation of these two main purposes of evaluation as set out in the opening paragraphs of UNDP’s new draft policy for evaluation quoted below, with the addition of a suggested change as highlighted. This would both better reflect the Norm, and give more substance to the accountability function.

4. “Evaluation ensures UNDP’s accountability for the achievement of its objectives through assessing its programmes and operations, including advocacy, advisory services, knowledge networks, technical assistance, coordination and partnerships. UNDP’s results will be evaluated for their *expected and achieved* contribution to the development effectiveness of partner countries.
5. Evaluation in UNDP will facilitate learning and improved performance by providing objective feedback of lessons learned from development experience. Evaluations will address what works and why, as well as what does not work and unintended outcomes. This will support informed decision-making and UNDP to position itself to better address development challenges.”

*) *This is different from accountability for the use of public funds in financial and legal terms, usually the responsibility of auditors and legal specialists.*

making the best use of the information to hand, therefore lack a basis for rigorously documenting achievements against stated and benchmarked objectives, and tracing linkages between activities, outputs, results and impact. They have value as an informed, professional and independent judgement, but fall short of an evidence-based account of UNDP's impact and effectiveness.

72. Related to the issues within the organisation's monitoring systems are questions of quality control of the outputs of the decentralised evaluation system. This should be a strength of UNDP's evaluation system, but reported experience is that the coverage and quality of evidence presented in decentralised evaluations are insufficient to underpin the programme of central evaluations completed or envisaged for the future. As importantly, given the substantial resources actually devoted to decentralised evaluation and planned for the future, the need to use such funds more cost-effectively is pressing.
73. We acknowledge the current practical difficulties in attempting to evaluate against intended results. However, we also conclude that the draft Evaluation Policy, while including several references, does not set its sights firmly enough on evaluating against intended results. As it stands, this sets accountability as a relatively low priority for the UNDP's evaluation system when compared with most organisations in the field, and places much greater emphasis on lesson learning alone.

Suggestions

74. We would suggest that performance benchmarks for each evaluation should be more explicitly discussed, and if possible agreed, with the relevant operational staff during the evaluation process. When developing benchmarks, the Panel would also note that the EO might wish to consider at least some level of benchmarking performance against that of other agencies, not least by reviewing relevant evaluative evidence from such agencies, if lesson learning is a major anticipated output of its evaluations.
75. We also suggest the new Evaluation Policy more explicitly emphasise the objective of evaluating against intended results, so taking this important opportunity to strengthen UNDP's evaluation culture and methodology in this key respect. The Panel is unanimous that the EO should always attempt to include explicit and transparent evaluation against the intended results of each undertaking, as far as practicable. We hold differing views however on whether or not the EO should also seek to rate UNDP performance in its evaluations, as is done by some peers. There is the view on the Panel that this should be an aim of the EO, since it would allow a richer degree of analysis in future Development Effectiveness Reports. Other Panel members however, are unconvinced that it is technically feasible to develop a performance rating system that is both soundly evidence-based and produces an accurate and fair reflection of actual performance.
76. In order to get better value from the UNDP evaluation system, EO needs to have a carefully defined role in both quality assurance and quality control of decentralised evaluations. Quality assurance measures refer to those "upstream" process steps that can be taken to provide a strong base and methodology for evaluation, while quality control refers to the "downstream" systems for testing the quality of evaluations produced. Probably regional-based evaluation expertise, as a link between country

level evaluations and EO, and a resource centre for decentralised evaluations, are part of the answer. Some mandatory procedures are also probably needed – e.g. for the regional expert to agree the terms of reference and methodology for each decentralised evaluation and review the evaluation report, plus regular interchange (and probably a reporting relationship) between the regional expert and EO. Further references to the needs, plans and options for these quality functions are included in Paragraphs 95, 109ff. and 178 of this report, among others.

Competence and Capacity

Is the professional competence and capacity of the Director and staff to deliver credible evaluations assured? (Norms 2.5 & 9.1-9.3).

77. The job descriptions and recruitment qualifications for the Director and professional evaluation staff of the EO include criteria for the appropriate technical and managerial competencies and experience expected. These criteria are applied during the selection process, which follows standard UNDP recruitment procedures. It is important that the Director of Evaluation have the authority to recruit internally and externally as necessary. Staff entering the EO from other positions within UNDP will rarely have first-hand evaluation experience, but relevant competencies are sought, and the Director of Evaluation has the final judgement in determining technical competency requirements. Developing the evaluation skills of those rotating in from elsewhere in the organisation relies heavily on intensive internal peer support within the EO. EO staff recruited from outside of UNDP are expected to have proven evaluation skills. The EO also uses consultants on fixed term contracts to augment its capacity in headquarters.
78. Competence and performance are then addressed during the annual Results Competency Assessment (RCA), in which an individual's performance (including completion of an agreed learning plan to fill specific knowledge gaps) is assessed by their line manager in accordance with UNDP-wide criteria.
79. As the major part of evaluation work is carried out by consultants, credibility also relies on the competence of the Evaluation Team Leaders and their teams. In the five evaluation processes that we examined in some detail, three of the Team Leaders appear to have been competent and capable across the range of tasks required. In the other two cases, the Evaluation Managers and other team members reportedly worked to complement the competence of Team Leaders in some areas that would normally be expected in the Team Leader.
80. The EO acknowledges that it needs to develop a larger roster of competent Team Leaders as the number of evaluations expands. It is thereafter considering adoption of the e-based roster system being tested and used by several other UNDP HQ units and regional offices. This would address the issues of availability and reducing the potential for conflicts of interest by providing access to an expanded body of consultants identified by UNDP's Bureau for Development Policy and others. However, some have suggested that it will be difficult to expand the roster of high quality consultants, since UNDP fee rates are currently considered low and obtaining a special waiver to pay competitive rates is a bureaucratic process. The EO also recognises the need to improve its forward planning to ensure that it can approach

and contract suitable Team Leaders in good time and before they become committed to other work.

81. Surprising to Panel members was the finding that while the EO strives to include evaluation competence within a contracted team, practice suggests that evaluation expertise is not always a vital criterion used in selecting Team Leaders. The record shows that an intimate understanding of UNDP and/or the reputation and stature to engage with high level partners are valued more highly in the selection of Team Leaders. In at least one case, this deficiency of evaluation expertise by the Team Leader was not even compensated for elsewhere on the team, except by the Evaluation Manager.
82. Our examination suggests that the Evaluation Office's own capacity to manage evaluations is stretched, and will be much more so with the planned expansion of the evaluation programme. This strain on capacity is partly because Evaluation Managers directly participate in most EO evaluations more intensively than their counterparts in many development agencies. Managers are expected to:
 - (i) Carry out a thorough review of evidence and consult widely when preparing the evaluation ToRs and recruiting and orienting the team;
 - (ii) Play a role as a *de facto* team member in country programme evaluations;
 - (iii) Devote significant time to both ensuring that stakeholders are consulted on the draft report;
 - (iv) In several cases, then devote significant time to redrafting the recommendations and ensuring that they are cast in the language used within UNDP; and
 - (v) Ensure that the EO as a whole judges that the evaluation product is of sufficient quality to be published in the name of the EO, which is the sole author of all reports.

Conclusions

83. Interviews with a wide range of stakeholders, including consultant evaluators and operational staff, confirmed a general perception that the staff of the EO comprises a professionally competent team. In only one evaluation (not a reference case) was the requisite competence of a contracted staff member of the EO questioned. We also conclude that the Office musters the requisite competence at the managerial level.
84. We concluded that a key determinant of EO's overall capacity is a corps of Evaluation Managers combining evaluation expertise, managerial skills and business and institutional knowledge. These Managers may be managing several evaluations at any one time. Given the range and intensity of their current tasks, we concluded that this places extraordinarily onerous demands upon the managers. Moreover, the shifting nature of their roles as now defined in the course of an evaluation (from organiser and recruiter, to team member, to quality controller) may reduce the transparency of the process and the credibility of the final evaluation product.

Suggestions

85. In the light of comparable experience, we believe that some redefinition of the tasks of Evaluation Managers may be necessary in order to be able to recruit, train and retain enough professional staff resources (especially for an expanding evaluation

programme) and also to clarify some of the anomalies among their current tasks and roles. Other systems with which we are familiar have various ways of organizing the allocation of roles and responsibilities of evaluation managers in particular evaluations. These may range from serving as organisers, resource persons and initial quality controllers, all the way to serving as fully fledged Team Leaders. We see definite potential for the Evaluation Office to streamline its own system.

Evaluability

Does the EO verify if there is clarity in the intent of the subject to be evaluated, sufficient measurable indicators, assessable reliable information sources and no major factor hindering an impartial evaluation process? (N7.1 & 7.2).

86. Resources available and time schedules dictate that EO managed evaluations must rely and build to a large extent on information that is already available. In terms of identifying UNDP's results and contribution, this means relying mainly upon information available through the RBM systems. Ensuring evaluability is therefore importantly a management function. UNDP's Programming Manual and Handbook of Monitoring and Evaluation (2002) both stresses the importance of operations staff setting clear objectives, indicators, baselines and targets, and ensuring collection of the necessary performance data. Evaluations commissioned by operational units tend to have smaller budgets, and thus pre-assessments of evaluability rarely take place.
87. Basic evaluability can be set up at the planning stage of activities, through setting clear objectives and clear indicators to measure progress and success. It can then be enhanced by proper monitoring of the implementation process as well as evaluation of individual projects/programmes at the output and outcome level. The former is clearly not a task for the EO, while the latter may be enhanced by the EO providing guidance and quality control (decentralised evaluations). Finally, the EO can conduct an evaluability study in order to be able to scope the evaluation and work out an evaluation plan.
88. Given the often weak foundations for evaluability in UNDP, two approaches to addressing this issue have been used for EO evaluations.
89. For country programme evaluations, evaluation managers have undertaken an initial scoping mission, during which evaluability has been assessed to some degree. However, the EO acknowledges that evaluability is a key concern for its country programme evaluations and that its major response to date on enhancing evaluability has failed. To make up for a lack of data from decentralised evaluations and the RBM system, local consultants have been commissioned in all country programme evaluations to gather and synthesise data in order to fill in gaps identified. Both consultant evaluators and the EO itself acknowledge that these commissioned studies have not significantly enhanced evaluability, except in one case, although the reasons for the shortcomings may sometimes include poor definition of the tasks as much as poor work by the consultants.
90. In 2005, the EO has formalised assessment of evaluability in its most recent country programme evaluation (in Serbia-Montenegro) by making this the main focus of the initial mission, with the aim of determining: (a) whether the Country Programme is evaluable; and (b) helping structure the evaluation approach accordingly.

91. For thematic evaluations, the evidence suggests that there is no consistent approach used within the EO to assess and enhance evaluability of individual evaluations. For the six thematic evaluations completed between 2002 and January 2006, Approach Papers were prepared in two cases by the contracted Evaluation Team Leader (and team) and were then agreed with the EO. While there is no evidence that these Approach Papers made a substantial impact on the conduct of the evaluations in these cases or significantly enhanced their evaluability, it remains good practice and helpful for an evaluation team to produce an approach paper that will clarify for all concerned how the basic Terms of Reference will be carried out.

Conclusions

92. Our major conclusion is that the evaluability has been severely and consistently constrained by the performance within UNDP's RBM systems, as illustrated in the below quotes from two EO publications from 2005:

“Establishing a viable frame of reference for measuring results in terms of outcomes and impact of the Regional Cooperation Framework and its constituent programmes has proved to be perhaps the most significant challenge in undertaking this evaluation, particularly as data was not being collected and monitored for the purpose of assessing outcomes and impact.”^{xv}

“Results orientation is weak in UNDP's programming process in Jamaica. Systematic monitoring and evaluation have been limited, and as a result, the process of lesson learning has suffered.”^{xvi}

93. For decentralised evaluations carried out under the responsibility of operational units, the proposed earmarking of budget shares for evaluations at the project and outcome levels is a positive step. In our view, this needs to be accompanied by stronger, and more proactive, backup in terms of evaluation advice, right from the planning stage, and at each significant milestone. This advisory capacity needs to be reinforced at either country or regional office level and therefore we endorse the efforts by the EO to co-fund, with the regional bureaux, evaluation expertise in UNDP's Regional Centres. We also conclude that the integration of updated evaluation guidance in UNDP's Programming Manual is a worthwhile step, since we found the relevance and quality of the 2002 monitoring and evaluation guidance to be still high. The costs of putting in place these measures would therefore be a sound and cost-effective investment, given the level of resources already devoted to decentralised evaluation.

Suggestions

94. Addressing the lack of data within the RBM systems is not directly within the control of the EO, but relies upon whether or not UNDP's Senior Management can successfully address the problems apparent within the RBM and decentralised evaluation systems.
95. For decentralised evaluations, the Panel therefore endorses the approach outlined in the draft Evaluation Policy, including the requirement that operational programmes earmark some agreed minimum level of programme funds to carry out planned evaluations. We also support the proposed contribution by the Evaluation Office to include supporting quality assurance through defining appropriate standards and

capacity building inputs for UNDP staff and would suggest the need for the Office (or its regional representatives) to have clear mechanisms for quality control of decentralised evaluations. We also endorse the introduction of evaluability assessments, but also note that these should not be applied dogmatically as this might preclude any evaluation.

96. However, we also conclude that the availability and quality of data available for EO evaluations are unlikely to increase substantially until well into UNDP's next corporate strategy framework cycle (2008-2011), since this depends on renewed effort in defining useable and relevant programme objectives and indicators across all programmes and this work will realistically start with introduction of the new Multi-Year Funding Framework, in 2008.
97. While EO's experience with commissioning local consultants to produce preparatory or foundational studies has reportedly been disappointing, we would urge against abandoning it. Properly defined and staffed, these tasks can play a vital role in grounding evaluations, as well as strengthening the elements of country ownership and capacity building which have not figured very prominently in UNDP's evaluation practice to date.

Are the EO's Evaluations Impartial?

Are systems and approaches in place that ensure the impartiality of EO evaluations? (N5.1-5.3 & N 11).

98. Impartiality is the absence of bias in due process, methodological rigor, as well as consideration and presentation of both achievements and challenges. It also implies that the "rules of the game" for evaluations are set and applied transparently and the views of all stakeholders seriously taken into account (discussed further in Chapter 4 of this report). In the event that interested parties have different views, they must have a fair opportunity to contest and persuade, and failing persuasion, to have persistent differences reflected in the evaluation analysis and reporting.
99. EO-managed evaluations are described as products of the EO and the EO's Evaluation Managers apparently play a significant role in the drafting of many evaluations. However, neither the evaluation guidance produced within UNDP nor evaluation ToRs explicitly describe the respective roles of consultants and evaluation managers in completion of draft reports. The Evaluation Managers therefore play an anomalous role by international standards of both being a member of the evaluation team and also responsible for quality control, while the extent of the independent contribution by those outside evaluation teams is also murky.
100. The evidence shows that the EO places a very high importance on using consultants with either a detailed understanding of UNDP or of the issue being evaluated. Practically, this means that many of the key consultants used have worked for, or within, UNDP in the past. The most obvious source of potential bias therefore would be from consultants evaluating operational programmes with which they had had a direct previous engagement. The EO states that they always check for such potential conflicts of interest. We are satisfied that the evidence, in the main, supports the EO's statement, although in one evaluation examined there was a convic-

tion among operational staff that two senior members of the evaluation team clearly came in with set institutional preferences on some of the major organisational questions examined.

101. However, while there are obvious reasons for giving importance to strong institutional knowledge in any organisation's evaluations, this needs to be balanced against the benefits of having other forms of experience (especially in evaluation itself) within the team and the dangers that a team will be perceived to be evaluating against preexisting opinions or biases. We also believe that the use of consultants that are very familiar with UNDP, and its terminology and outlook, may be contributing to an external communication problem observed by all of the Panel members. While evaluation reports are generally well written, they often fall into a UNDP house style, which is difficult to penetrate from one reading of a report by readers not intimately familiar with UNDP.
102. While the EO invests significantly in consultation with operations staff and other relevant stakeholders during the evaluation process, we find that stakeholders' roles and responsibilities in this process are either not entirely agreed or understood by all stakeholders.
103. Finally, we could find no instances where evaluation reports actually included dissenting conclusions or opinions, although our work shows such differences do exist. We have concluded that in the genuinely open culture for debate within UNDP, not reporting dissenting views in the evaluation report increases the danger that an evaluation will be seen as biased. The move to institutionalize and simultaneously publish management responses for all EO managed evaluations will help manage this danger to some extent, but not entirely.

Conclusions

104. As a general rule, the evidence suggests that EO evaluations are carried out in an impartial manner, but we conclude that some practices may contribute to perceptions that the evaluations are not impartial. Wherever there may be such a risk, we are convinced that the best protection is to build in transparency as systematically as possible.

Suggestions

105. We suggest that the roles of key stakeholders in the evaluation process need to be set out very clearly in the Evaluation Policy and in the guiding documents for each evaluation, as a way of both increasing the credibility and transparency of the process, and also helping reduce instances of drawn-out debates significantly delaying completion of an evaluation.
106. We would further suggest that, in selecting consultants for future evaluation teams, UNDP's Evaluation Office give greater weight to evaluation experience and skills (and knowledge of the subject or theme to be evaluated) and lesser emphasis to "insider" knowledge of the institution. These requirements need to be covered in the overall strengths of a team, with special attention to the particular qualities required of Team Leaders in providing professional authority combined with high skills in team-management and writing.

107. Future evaluation ToRs and the Introductions of evaluations should be much clearer in defining the role of the consultant Team Leader and the Evaluation Manager/EO in the evaluation process as a whole and in drafting of the final evaluation product. The fact that the Evaluation Office takes full responsibility, and often an active editorial role, in the contents of evaluation reports should be made unambiguously clear, both as these evaluations are organised and carried out, and as their results are presented.
108. The simultaneous publication of management responses to evaluations and the explicit reflection of significant dissenting views in evaluation reports would also enhance the sense of impartiality and constructive debate around future evaluations. The new Policy should address the possibility of evaluation reports specifically reflecting dissenting views where warranted.

Assuring Quality

Does the evaluation process used by the EO enhance the quality of the evaluations produced? (N8.1 & N8.2).

109. It is only recently that evaluators in either the multilateral or bilateral development organisations have begun to develop rigorous standards for assessing the quality of evaluations. This means that both the definitions of quality, and views on which systems are most important for ensuring it, have varied among peers.
110. To date, as in many other organisations, quality assurance criteria in UNDP have been more implicit, and built into the judgement and oversight of evaluation managers, their peers, and senior management in the Evaluation Office. From our examination of the evaluations over recent years, this system seems to have functioned reasonably well. The UNDP's EO has, however, been working on developing and testing a set of quality criteria for assessing evaluations, which it is currently codifying into a final set of evaluation quality standards. It plans to introduce both the standards and institutional mechanisms for supporting and utilizing them by February 2006.
111. In terms of the evaluation process, we have already noted a number of issues that potentially affect the quality and credibility of the evaluation process. These include the lack of clarity in the roles of evaluation managers and consultant evaluators and of stakeholders consulted; the potentially onerous demands placed on the evaluation managers; and the relative low priority sometimes attached to evaluation skills when selecting team members.
112. In terms of the actual evaluation reports, we found that it is sometimes difficult to follow the links between evaluation findings, conclusions and recommendations. We also found that the EO's evaluation reports consistently included a large number of unprioritised recommendations of varying relevance to different stakeholder groups.

Suggestions

113. Given the intrinsic difficulties in maintaining and enhancing the quality of evaluations – and some of the important constraints still to be managed in the UNDP system – we would offer a word of caution about the ambitious plans to expand rapidly the number of EO-evaluations to be undertaken. As evaluators, we would counsel against diluting rigour and quality, especially of country programme evaluations, in order to increase coverage. This is a particular issue given the comparative lack of robust evidence available within UNDP's systems upon which the EO's evaluations can build. The "Assessments of Development Results" in particular are intended to be major evaluations of country programmes, and we would suggest that the aim should be to produce an appropriate number of well-selected and solid ADRs, with country offices and bureaux conducting other types of reviews to meet their management requirements across the remaining country programmes.
114. In terms of the evaluation process, we would suggest that quality might be further enhanced if the EO were to make greater use of internal, external and mixed expert panels and reference groups as "sounding boards" at key milestones in all of its evaluations, rather than in just selected thematic evaluations, as is present practice. This would help address the challenge of evaluating actual results achieved in an environment where results frameworks and indicators are often poor and where it is both theoretically and methodologically difficult to assert what the impact of UNDP's contribution to development results observed may actually be.
115. Credibility would be further increased by making the links between analysis, findings and recommendations in evaluation reports as explicit as possible. This can be done, for example, by cross-referencing to the specific finding when making a recommendation in an evaluation report.
116. Experience elsewhere also suggests that evaluations which feature a limited number of prioritized recommendations are more effective. More minor recommendations, in turn, can be better included in an administrative annex.

Summary of Conclusions and Key Suggestions

117. Our overall conclusion is that the credibility of evaluations produced by the EO is acceptable and improving in relation to international practice in this area when assessed against the relevant UNEG Norms, and could be further strengthened by a number of relatively straightforward measures. The Director, staff and consultants used are professionally competent, while the processes used to manage evaluations process is basically conducive to quality. While some perceptions of bias within evaluations were identified in the course of our discussions, we do not conclude that there is sufficient evidence to suggest that the system as a whole is vulnerable to significant bias.
118. We strongly endorse:
- The draft Evaluation Policy's intention to mandate a built-in allocation for decentralised project and outcome evaluations in operational budgets, which

(if conducted in an adequate fashion) will help provide a more solid basis for higher-level evaluations in future;

- The decision of the EO to introduce “evaluability assessments” as a normal preparatory step in their evaluations;
- The intention to introduce a set of evaluation quality standards and institutional mechanisms for supporting and utilising these from early 2006, with a carefully defined role for the EO in both quality assurance and quality control of decentralised evaluations; and
- The agreement to strengthen and systematise the production, and simultaneous presentation, of management responses for all evaluations presented to the Executive Board. The new Policy should address the possibility of evaluation reports specifically reflecting dissenting views where warranted.

119. These initiatives hold the potential for addressing some of the weaknesses in the present evaluation approach. We have also offered a number of other suggestions aimed at increasing transparency and credibility by:

- Strengthening the elements of performance assessment and accountability – particularly the objective of evaluating against intended results – in the Evaluation Policy and individual evaluations (while flagging again the problems resulting from the documented weaknesses of UNDP’s systems of managing for results);
- Clarifying the “rules of the game” for all stakeholders in the Evaluation Policy and in guiding documents for each evaluation. For example, the fact that the Evaluation Office takes full responsibility, and often an active editorial role, in the contents of evaluation reports should be made unambiguously clear; and
- Taking steps to reduce occasional perceptions of a lack of impartiality.

120. Our examination of evaluation practice in UNDP has revealed relatively little emphasis to date on the involvement and “ownership” of partnercountry stakeholders in these evaluations beyond serving as interviewees, and participants in follow-up workshops in the cases where these are held. The UN Norms for Evaluation do not address this as a major issue while the present draft Evaluation Policy for UNDP sets out more ambitious principles, for example through more joint evaluations and evaluation capacity building. While the nature of some of UNDP’s thematic evaluations may not be especially conducive to intensive engagement by programme country stakeholders, we suggest that the Evaluation Office seek creative ways of strengthening country programme involvement and partnership in all its work and other UNDP evaluation activity. To make this a priority would be consistent with good practice and especially UNDP’s own vocation to support country ownership and capacity building. Some examples might include: formal joint evaluations of country programmes; regular inclusion of partnercountry experts and trainees on evaluation teams; and support for evaluation associations and professional development activities in the field.

121. Welcome early steps are now being taken to strengthen UN coordination around evaluation, and presumably the EO will be working to operationalise the other principles featured in the draft Policy around human development and human rights and ethics.

4. Utility

“To have an impact on decisionmaking, 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 of the many parties involved in development cooperation.” xvii

The Purposes, Importance and Focus of Evaluation

Are the purposes and importance of evaluation clearly accepted in UNDP? [N1.1]

Do evaluations examine the rationale and justification of undertakings and the satisfaction of intended beneficiaries, effectiveness and efficiency in achieving expected results, alternative approaches and lessons? [N1.7].

122. As would be the case in many other organisations today, the Panel found in UNDP a partial and uneven understanding and acceptance of the purposes of evaluation, and of its importance. At the top governing level, that of the Executive Board, the awareness of this instrument and its uses is clearly not strong, although it is beginning to be called upon more frequently, at the initiative of a number of members from all regions. Overall, however, the Board is still under-exploiting the potential of the evaluation to serve as a “reality check” on the organisation’s self-assessments of results under its Multi-Year Funding Framework and reports on annual results. At the different levels of management, there is a definite acceptance in principle of the legitimacy of evaluation, especially for purposes of learning, but the systems for management engagement are not strong enough to give this instrument its appropriate importance among the various streams of management information. The potential of evaluation as a key tool of accountability for assessing performance of the organisation’s work is much less clear or accepted, although most respondents on the Panel’s reference cases saw at least some accountability function being served.
123. With respect to the level of focus, the overall finding is that the most important single focus of most recent evaluations has been the attempt to identify clearly results achieved, and UNDP’s contribution to them, as a base for future directions. Terms of Reference usually call for an examination of the rationale and justification of an undertaking, and we find that this has been done to varying degrees in the actual evaluations; with the Global Cooperation Framework and country programme evaluations probably doing this most successfully. Given the nature of UNDP’s activities, the satisfaction of intended beneficiaries has been a lesser focus, although the satisfaction of differing partners can be an important issue in evaluations, especially where some evaluation findings have been hotly contested, and where UNDP’s own surveys of partner and staff satisfaction portray a different picture.
124. It has proved even more difficult for EO evaluations to examine the effectiveness and efficiency of achieving expected UNDP results. This is primarily because the expected results have not been sufficiently clearly established by operational units at the planning stage, and/or because performance against them has not been monitored and evaluated at the level of projects and outcomes. In turn EO attempts to strengthen evaluability ex post have not been effective.

125. For the country level Assessments of Development Results, as they now stand, the approved evaluation framework and guidance actually specifies that they are expected to: “Combine a ‘goal-free’ evaluation with assessing specific goals. Based on an assessment of key results and past achievements (or failures) in the areas UNDP has supported over the last five years or so, the ADR will provide a forward-looking analysis.” The implication of applying a “goal-free” evaluation is that there are no preset benchmarks for these evaluations, setting major limitations on their utility both for accountability and learning.
126. Bearing in mind the limitations just noted, UNDP’s centrally managed evaluations have given a considerable amount of attention to examining alternative approaches and lessons learned and have proved useful in pointing out some new directions and/or in informing careful thought and debate about existing directions. This was consistently identified as the major benefit of the ADRs by operational managers.

Conclusions

127. We conclude that the base of understanding of evaluation and its importance is reasonably strong in UNDP, and that it has been getting somewhat stronger in recent years, including some greater interest at the Executive Board level. At the same time, we conclude that only some of the purposes of evaluation are being properly pursued. In particular, the potential of evaluation as a key tool of accountability for assessing performance of the organisation’s work is much less clear or accepted than that of learning. This is particularly the case for country programme evaluations in their current form. Most respondents on the Panel’s reference cases saw at least a potential, if not yet substantive role, for evaluation in meeting UNDP’s accountability requirements.
128. Against the benchmark of the Norms, the rigor of the EO’s evaluations is clearly impaired by the lack of hard quantitative evidence, and the need to rely excessively on qualitative judgements and benchmarks defined by the evaluators’ themselves. This limits their potential usefulness for accountability. However, the Panel would also state that a similar finding would be found with any evaluation office assessed, which illustrates the challenge of creating effective decentralised evaluation and RBM systems across the development community.

Suggestions

129. The discussion and approval of a new Evaluation Policy should serve as an important opportunity for further clarifying and agreeing the purpose of evaluation, and we make a number of suggestions elsewhere that might further assist in bolstering its use. Overall, we strongly suggest that the Evaluation Office reexamine and adjust the current heavy emphasis on learning from evaluation, to give greater weight to accountability.

Contributions to Managing for Results

Does evaluation make an essential contribution to managing for results, and inform the business cycle? Does it aim at improving relevance, results, the use of resources, client satisfaction and maximizing the contribution of the UN system? [N 1.3].

130. As noted earlier, the Panel has found that UNDP not only shares the difficulties faced by all development cooperation agencies (and others) in putting in place effective systems to manage for results, but that some of these difficulties may be particularly severe in this organisation, in part because of the increasingly “soft” character of much of its work, as the organisation has tried to respond to the challenge of becoming a supplier of ‘up-stream’ policy advice to partner governments.
131. It appears that for a period, the effort invested in developing an RBM system (spearheaded at the time by the Evaluation Office) actually displaced earlier evaluation work and roles. These are now being re-established and the evidence suggests that the EO’s evaluations are generally targeted and timed to affect specific decisions made at key points in the programming cycle.
132. Given the pivotal position of UNDP within the UN development family, the roles of Resident Coordinators and the UNDP’s Evaluation Office among its UN peers, there is a definite concern to help maximize the overall contribution of the UN system, for example through the work of the UN Evaluation Group, and the plan to develop evaluations of the concerted UN Development Assistance Frameworks (UNDAFs) within countries.

Conclusions

133. In spite of the limitations noted, the Panel has found that recent UNDP evaluations are geared and timed to decisionmaking and have a growing potential to contribute more to the business process cycle, with increased management “take-up”. The trend is upward, and new opportunities are clearly emerging, both in thematic and country level evaluations.
134. The extensive awareness and documentation of the weaknesses in UNDP’s current systems for managing for results should lead to substantial improvements by management, and presumably the work currently underway on a new accountability framework will help the organisation move in that direction. In that case, evaluations should be in a position both to support moves to a better functioning RBM system, while also benefiting at the same time from the stronger foundation and “building-blocks”.

The Intention to Use Evaluations

Is there a clear intention to use evaluation findings and are evaluations planned and targeted to inform decisionmaking with relevant and timely information? [N 4.1].

135. The Panel has found that when evaluations are undertaken in UNDP there is a strong intention to use their results, particularly in relation to the preparation of future programmes. In relation to thematic and regional evaluations, this connection has recently been reinforced by the explicit and serious linkage in Executive Board deliberations of evaluations (those of the Second Global Cooperation Framework and of the Regional Cooperation Framework for the Arab States) with decisions on the succeeding programme frameworks.
136. In four of five reference cases examined, the panel found that the planning, scheduling and delivery of the evaluations was relatively well targeted to decisionmaking

milestones, whereas the fifth case was an anomalous situation where the milestones themselves were never really clear. The relevance of the information provided by the evaluations was unquestionably very high in two of the four targeted cases, and was considered high by key participants in the other two, even while some of the conclusions themselves were vigorously contested by some parties.

Conclusions

137. On the key aspects of an intention to use evaluation results and timely planning and targeting to make that possible, we have found a comparatively close meshing of demand and supply for evaluation in UNDP, and a potential to take this link to a higher-level. Under somewhat difficult circumstances for planning and budgeting, the Panel has found that the Evaluation Office has been able in recent years to plan and maintain a rolling work programme of centrally managed evaluations with substantial relevance and broad coverage of thematic and regional issues. Meanwhile, the approach to country programme evaluations (ADRs) has been explicitly one of experimentation to develop the most effective approach.

Suggestions

138. We would strongly suggest that after the past three years of experimentation with the current approach to country programme evaluations, a definitive model should be adopted as soon as possible, rectifying as far as possible the limitations in the current “goal-free” evaluation approach and supported by a strong methodology.

Transparency and Consultation

Is there transparency and consultation with the major stakeholders in all stages of the evaluation process, and are the key documents made available in readable form?

[N 10.1 & N 10.2].

139. The evaluation process in UNDP appears to allow (and to encourage in the Guidance documents) full transparency and consultation with agency managers and staff, from developing the ToRs through to development of the recommendations. Practice is variable, and is rarely laid out in a coherent consultation plan, in which the time period for consultation at each step has been clearly defined and agreed between the parties involved. Thus in each of the reference cases for this assessment there was a different level of awareness and engagement in consultation between evaluation managers, consultant team members and leaders, and operational personnel involved. The experience suggests that for evaluations not facing a hard completion date for presentation to the Executive Board, this may lead to drawn out and potentially unproductive consultation processes. This was the case in one Thematic Evaluation, where, by the time the evaluation was completed, events had rendered its value minimal, and has reportedly been the case in at least two ADRs. Reportedly, the engagement of stakeholders in decentralised evaluations has also been variable in the planning/design stages and follow-up of evaluations.
140. The evidence also suggests that roles and responsibilities of differing parties in the consultation process have not always been either understood or agreed. In particular, the principle that consultation on draft reports should focus on whether or not

the factual evidence and findings are correct but that the conclusions and recommendations are not the subject for consultation has been difficult to implement. It would appear that this principle has been difficult to implement for two reasons. First, because the EO does not record dissenting views in the evaluation reports. Second, because the system of using management responses has only recently been introduced.

141. Final reports of centrally managed evaluations are made readily available to all in both printed and electronic form. The terms of reference are included in all reports. Linked to a lesser investment in dissemination, however, the Panel encountered several instances around its reference evaluations where some participants were left unaware for long periods of the status or even the publication of reports. In terms of being in easily consultable and readable form, the Panel found that the Evaluation Office invests considerable effort in meeting this perennial challenge, and relatively successfully from the cases examined, although there is a tendency to target the language of reports to an “inside” audience.

Conclusions

142. We are sensitive to the fact that achieving the necessary degree of transparency and consultation with all stakeholders at all stages of evaluations is usually very demanding. For a start, it requires synchronized involvement of the evaluation office, teams and the managers and other stakeholders concerned at the outset and at key milestones along the way. Second, for UNDP’s far-flung partnership programmes in particular, it may be very difficult to push very far down the ladder toward beneficiaries. Third, it is common that some of the stakeholders may take little interest at an early stage when they may see the exercise a somewhat abstract and distant, but become intensely interested later when findings and conclusions begin to emerge. Failing this consultation, however, at this point the base for ownership may not have been built, the take-up of results is likely to be weaker, and indeed the due process and good faith in the process can be called into question.

Suggestions

143. We suggest that further clarifying mutual expectations among parties on the subject of consultation in the new Evaluation Policy as well as drawing up and following a full roadmap for consultation in each evaluation, clearly specifying these arrangements, will be more than worth the additional effort.

Management Response and Follow-up

Are evaluations followed up with explicit responses by those addressed by their recommendations, systematic tracking and periodic reporting of the action taken on accepted recommendations? [N 12.1-12.3].

144. At present, the Panel found all three of these follow-up systems to be very weak, although some recent steps have already been taken in the right directions, and more substantial ones are planned. Since 2003, management responses have been produced for four of the five evaluations presented to the Executive Board. However, no regular system is yet in place for systematic and regular reporting on

implementation of agreed actions in response to recommendations. The management response to the GCF II (Global Cooperation Framework) evaluation in 2004 was apparently the first to lay out a fairly detailed action plan as described, although the further follow-up and monitoring arrangements were not clear. The new Regional Cooperation Strategy for the Arab States was presented in conjunction with the evaluation of the previous Strategy and drew extensively upon it, with due acknowledgement. However, no separate management response was produced.

145. To date, country programme evaluations, or Assessments of Development Results, have not required a formal management response. The Evaluation Office confirmed that they have just introduced such a “Management Response Table” laying out graphically:

- Key Recommendations;
- UNDP/UN Unit responsible;
- Management Response to recommendations; and
- Strategy for addressing issue, including indicators of progress, partner ships and timeframe.

146. To date, the Guidance for such Assessments of Development Results has envisaged that a follow-up workshop in-country would be the most important single follow-up measure for these evaluations. In practice, however, for only three of fifteen such evaluations conducted have such workshops been held, with all three workshops being reported to have been quite useful.

Conclusions

147. To ensure the effective use of evaluation in UNDP – while also reinforcing its credibility and independence – it is hard to overstate the central importance of improved arrangements for management response and the systematic follow-up of evaluation recommendations. Building on the precedents of the GCF II and Arab Regional Cooperation Framework evaluations, the draft Evaluation Policy offers the prospect of a much more rigorous system.

Suggestions

148. We suggest that the basic practice of setting out main findings, recommendations, unit(s) responsible, management responses, and time bound strategies for addressing them should be applied to all evaluations carried out in UNDP. In relevant cases, inviting and reflecting the response of partner countries would also be good practice. Follow-up action then needs to be monitored by management (perhaps with the support of an audit function) and reported by, and through, the Evaluation Office. This would then provide some capacity for monitoring the utility of evaluation in the organisation.

Knowledge, Evidence and Added Value

Does evaluation provide decisionmakers with knowledge and evidence about performance and good practices and add value for improvements in policy-making, developmental and organisational effectiveness? [N 1.5].

149. Evaluation is expected to provide one of several streams of performance management information in UNDP (together with Audit, RBM, and surveys of different kinds). Senior management acknowledges that these streams are not being sufficiently brought together, although there are also views that the RBM systems are providing information that should be a “rich empirical for the Evaluation Office’s work”. Near unanimity in the interviews for this assessment, the Evaluation Office’s own studies^{xviii} and a range of evaluation reports themselves all testify that the necessary foundations for rigorous evaluation to feed effectively into management and decisionmaking are still very weak (namely adequate results frameworks, together with the monitoring, evaluation and data building blocks needed).

Conclusions

150. While taking account of these limitations, the Panel has strong evidence that the central evaluation function is designed to provide added value in these areas. The evaluations produced are not shelved or treated as irrelevant or unpalatable. Combined with strategic scheduling, the choice of topics and the critical freedom adopted in evaluation processes and reports have ensured that they have stirred and informed debate and higher-order decisionmaking by management and in some cases the Executive Board, even in cases where some of the conclusions are contested. By comparative standards, this reflects a good use of evaluation. Prominent examples examined included the evaluations of the Arab Regional Cooperation and Global Cooperation Frameworks, and that on Gender Mainstreaming, with the last-mentioned having widespread impact even before its formal completion. The country programme reports examined showed smaller circles of impact, but were found to have served these purposes for the regional bureaux and/or country offices concerned.

Accessing and Disseminating Lessons

Is there is a repository of evaluations and a mechanism for distilling and disseminating lessons from evaluation, internally and externally? [N 2.7].

151. Each of these elements has received substantial attention by the Evaluation Office.
152. Since 2004, Evaluation Office products and outcome evaluations produced by country offices have been accessible to all UNDP staff through the Evaluation Resource Centre. The Evaluation Network (EVALNET) and the organisation’s Evaluation Forum are both considered to have been relatively successful mechanisms for disseminating evaluation findings. All EO managed evaluations are also available on its external website.
153. Fourteen digests of lessons learned – Essentials – were produced between 1999 and 2002. This instrument was discontinued for a period but is apparently being restarted. However, the evidence from the five reference cases examined is that the EO does not include funds within evaluation budgets either to synthesise the results of an evaluation or repackage the findings/recommendations for targeted audiences. Neither does it make use of either evaluation consultants or evaluation managers in dissemination.

154. However, as well as presenting selected evaluations to the Executive Board, the EO also drafts the Administrator's Annual Report on Evaluation, which does include some synthesis of findings and results. The Development Effectiveness Reports produced by the Evaluation Office, once every two or three years, also aim to synthesise information from evaluations relevant to the broad theme addressed in a particular DER.
155. With respect to the availability of findings, UNDP's disclosure policy includes a presumption in favour of public disclosure of all of the information and documentation generated or held by UNDP. Neither evaluation staff nor others have explicitly identified either accessibility or disclosure as a problem with centrally managed evaluations.
156. With a view to maximizing the actual use of evaluation findings – and at the same time reinforcing their credibility and independence – the Panel has found in UNDP, in common with most of their experience elsewhere, much lesser attention being paid to disseminating findings, relative to the huge efforts invested in producing the evaluations.

Conclusions

157. We conclude that current accessibility and disclosure arrangements are appropriate for the results of centrally managed evaluations. We have not been able to assess directly the provisions for accessibility and disclosure of “decentralised” evaluations, but on the basis of available indications we would question their adequacy.
158. By the time the extended process of an evaluation is completed, a natural tendency is to turn quickly to starting the next. We share the conclusion that in tandem with improvement of the quality and coverage of evaluations, a high priority for the Evaluation Office's attention and resources should be in possibilities for stimulating more active demand for readable and useable evaluation results, and proactive measures to generate openings for them. The Evaluation Office demonstrated that it is aware of this challenge, and is putting in new effort and actively studying successful dissemination practice elsewhere.

Suggestions

159. Building on the current practice for accessibility and public disclosure of the results of centrally managed evaluations, we suggest that similar arrangements should be extended to the results of decentralised evaluations. All should be put in the public domain, in the interests of quality, transparency and lesson-learning.
160. We suggest that the Evaluation Office invest energetically in the specialised attention and skills to succeed in proactively disseminating of its findings in user-friendly forms. Experience shows that this requires a “champion” armed with a strong mandate and creative skills for this work. For example, UNDP's Evaluation Office could strive to become a leader in this area by:
 - Building the time and money needed for dissemination tasks right into its approach papers and into the duties of task managers and the terms of reference for consultants;

- Developing a fuller range of dissemination tools, targeted at particular audiences; and
- Systematically seeking out and being alert for relevant openings or “hooks” for injecting its evaluation findings into processes or events organised by others. This may prove just as effective as direct dissemination efforts around its own products.

The Management of Evaluation

The evaluation plan can be the result of a cyclical or purposive selection of evaluation topics. The purpose, nature and scope of evaluation must be clear to evaluators and stakeholders. The plan for conducting each evaluation must ensure due process, a cost-effective approach, and the timely completion of the mandate. [N 4.2].

161. Overall planning. The biennial work programme for evaluation is available on the EO’s website, although the rationale and processes for its planning and the selection of the subjects for evaluation (e.g. cyclical or purposive) are not prominently explained, either on the website or in the Annual Report on Evaluation.
162. Individual evaluations. As noted in the earlier discussion of transparency and consultation, the Panel found, in the reference cases examined, that the purpose, nature and scope of evaluation have not always been clear enough to evaluators and stakeholders to ensure the most constructive process and useful results. In two of five cases – one country programme evaluation and one at the thematic level – this lack of clarity seems to have been a serious concern. In one other thematic evaluation, there was extensive engagement and debate which may never have been able to dispel differences on these issues because of the high and contentious institutional stakes involved. In the other two cases, the purpose, nature and scope of evaluation seemed relatively clear to all.
163. While this is an art as well as a science, because the factors are never entirely predictable, a plan for each evaluation can help greatly with these challenges and those of cost-effectiveness and timely delivery. It is also essential to ensuring due process, since there is no other instrument readily available to all concerned in an evaluation that will specify what steps are to be followed, by whom, when, and how. The Panel found that such plans (which are often called “inception reports” or “approach papers”) were probably sufficiently well laid out for participants in two or three of the five reference evaluations, while the others were less clear.
164. With respect to cost-effectiveness, the Panel gained the sense that, by comparative standards, individual evaluations, especially those of country programmes, may not be adequately funded, given the uneven base of existing information with which they must work. Some of this economy may be illusory, if short preparation time or other factors seriously impair the quality of the process or the product.

Suggestions

165. Publicly clarifying the process and the criteria applied for preparing the evaluation plan should be a useful step in regularizing expectations about evaluation in UNDP and at the same time enhancing its independence and credibility.

166. As noted in the earlier discussion of reporting to the Executive Board as a pillar of the independence of evaluation in UNDP (see especially Paragraph 20) the Panel sees an important need for the Evaluation Office to target carefully the future roles of the Annual Report on Evaluation and the periodic Reports on Development Effectiveness as key vehicles for communicating: (i) the EO's work in the past year and its plan for the coming year, together with an overview of "decentralised" evaluation activity; (ii) consolidated results and impact, including crosscutting issues relating to UNDP's performance as evidenced by evaluations; and (iii) a report that summarises whether management are implementing agreed evaluation recommendations on schedule. EO's strategic reporting on the evaluation programme should also serve to clarify and reinforce the logic of planning and coverage of the programme of centrally managed evaluations, and explain the process by which the plan is arrived at, as well as highlighting the main results and lessons of the evaluations carried out. This might help encourage the purposive use of evaluation by all concerned, and also help protect it if necessary.

Summary of Conclusions and Key Suggestions

167. We conclude that the base of understanding of evaluation and its importance is reasonably strong in UNDP, and that it has been getting stronger in recent years, including some greater interest at the Executive Board level, although the Board is still underusing the potential of evaluation to serve as a "reality check" on the organisation's self assessments of results. At the same time, we conclude that only some of the purposes of evaluation are being pursued, and in particular its accountability and performance assessment purposes are being underexploited, relative to those of learning. This is particularly the case for country programme evaluations (ADRs) in their current form. Overall, we strongly suggest that the Evaluation Office reexamine and adjust the current heavy emphasis on learning from evaluation, to give greater weight to accountability.
168. We would strongly suggest that after the past three years of experimentation with the current approach to country programme evaluations, a definitive model should be adopted as soon as possible, rectifying as far as possible the limitations in the current "goal-free" evaluation approach.
169. Recent UNDP evaluations are geared and timed to decisionmaking and have a growing potential to contribute more to the business process cycle, with increased management "take-up". The evaluations produced are not shelved or treated as irrelevant or unpalatable – they have stirred and informed debate and higher-order decisionmaking, even in cases where some of the conclusions remain contested. The trend is upward, and new opportunities are clearly emerging, both in thematic and country level evaluations.
170. Achieving the necessary degree of transparency and consultation with all stakeholders at all stages of evaluations is demanding, but without it the base for ownership may not have been built, the take-up of recommendations is likely to be weaker, and indeed the due process and good faith in the process can be called into question. We suggest further clarifying mutual expectations among parties on the subject of consultation in the new Evaluation Policy as well as drawing up and following a full road-map for consultation in each evaluation, clearly specifying these arrangements.

171. We judge that no areas are more important to ensuring the effective use of evaluation in UNDP – while also reinforcing its credibility and independence – than are improved arrangements for management response and the systematic follow-up of evaluation recommendations. Building on recent precedents, the draft Evaluation Policy offers the prospect of much more serious systems in these areas. We suggest that the basic practice of setting out recommendations, unit(s) responsible, management responses, and timebound strategies for addressing them should be applied to all evaluations carried out in UNDP, and that implementation should be reported upon through the Evaluation Office.
172. We conclude that current accessibility and disclosure arrangements are appropriate for the results of centrally managed evaluations. We suggest that “decentralised” evaluations should also be put in the public domain. We further suggest that the Evaluation Office invest energetically in the specialized attention and skills to succeed in proactively disseminating of its findings in user-friendly forms, and we make a number of specific suggestions to this end.
173. The Annual Report on Evaluation should serve to clarify and reinforce the logic of planning and coverage of the programme of centrally managed evaluations, and explain the process by which the plan is arrived at, as well as highlighting the main results and lessons of the evaluations carried out. It should also aim to capture highlights from the evaluations carried out in the decentralised evaluation system.

5. The Summative Judgement

174. The United Nations Development Programme has an Evaluation Office (EO) which enjoys an acceptable level of independence and which produces evaluations that are credible, valid and useful for learning and strategy formation in the organisation. At the same time, its potential contribution to strengthening accountability and performance assessment is being underexploited, both in its own right and as an essential basis for learning.
175. The Evaluation Office musters the requisite competences at the managerial and professional levels and has clearly been strengthening its role and performance in recent years. Building on this foundation, and using the new Norms for Evaluation in the UN System, the current testing of a new Policy for Evaluation in UNDP offers an extraordinary opportunity to further strengthen the function and its contribution to UNDP's results.
176. In this context we offer a number of suggestions and possible options for measures: first, to firmly secure its independence; second, to beef up the accountability and related performance assessment functions; and third, suggestions for more systematic stakeholder participation, management response and follow-up, and proactive, targeted communication of evaluation results.
177. While of necessity focusing on the central evaluation function, it is been important to take account of its place among the various information streams available for the governance and management of UNDP, and how it interacts with the others. In that light, we emphasise the need to reinforce UNDP's results based management systems, and in particular the quality and accessibility of information produced by these systems.
178. At the operational level, we endorse the EO's intention to support an increase in the quality and availability of "decentralised" evaluations (managed by country offices and bureaux) on the actual outcomes of UNDP's projects and other activities. In addition to their intrinsic importance for operational managers, these are the indispensable "building blocks" for a strategic evaluation function to serve properly the needs of the Organisation as a whole.

Annex I: A Description of Evaluation in UNDP

Oversight and Evaluation ¹

1. Evaluation is a long-standing and widespread function in UNDP, which is carried out by both the operational programmes of the organisation (decentralised evaluations) and by the Evaluation Office (EO). The EO is described as “an independent office reporting to the Executive Board through the Administrator” with the reporting line of the Director of the Evaluation Office being to the Administrator.”²
2. The other two key UNDP offices dealing with oversight are the Operations Support Group (OSG) and the Office of Audit and Performance Review (OAPR). OSG has primary responsibility for managing development of UNDP’s performance management and reporting system, principally through coordinating development of the Multi-Year Funding Framework (MYFF) and coordinating reporting against the framework to the Executive Board in the MYFF report. The EO works with OSG to try to ensure that evaluation remains aligned with the corporate planning framework found within the MYFF.
3. OAPR carries out internal audits within UNDP and all audit missions review ROAR production processes and quality. EO and OAPR currently share evaluation and audit plans and are looking at how to enhance production of more balanced assessments of country programme performance and UNDP’s contribution to development effectiveness. The introduction of mandatory evaluations of all country programmes proposed in the EO’s draft Evaluation Policy would be the main contribution from the EO in this area.

The EO’s Mandate

4. Current mandated activities for the EO include to:
 - Undertake strategic evaluations of UNDP management and programme policies, conduct assessment of specific organisational policy and evaluate the impact of UNDP funded programmes;
 - Promote the use of evaluation findings, lessons and recommendations in policy and programme formulation;
 - Ensure that evaluation findings are accessible to staff and stakeholders and are built into UNDP knowledge management system;
 - Develop evaluation guidelines and methods to assist programme managers in results management;
 - Develop methodological tools and systems to support results orientation in the organisation;
 - Maintain dialogue with programme countries and country offices to strengthen the internalization and utilization of evaluation standards both within UNDP and amongst its partners.

¹ UNDP Organigram can be found at www.undp.org/about_undp/2003organigram.pdf

² Response to Questionnaire for Self-assessment, UNEG Task Force on ‘Quality Stamp,’ 2005. Fact Sheet.

- Oversee the evaluation function in UNDP, including compliance with requirements and tracking of recommendations.
 - Prepare the Annual Report of the Administrator on Evaluation; and
 - Maintain and develop partnerships with the UN system organisations (as chair of UNEG), multilateral banks, OECD-DAC on evaluation related work to ensure infusion of globally acceptable quality standard of monitoring and evaluation, partner governments, national and international networks (African Evaluation Association; IDEAS etc).
5. At the June 2005 Executive Board meeting, the Board approved the codification of an evaluation policy for UNDP, to be presented for approval at the June 2006 Executive Board meeting. In September 2005, UNDP senior management endorsed the Policy and agreed that the EO could field test it prior to its presentation to the Board. The draft Evaluation Policy, as currently proposed, covers all of the issues identified in the UNEG Norms as necessary, as shown below:
- Explanation of the concept and role of evaluation within the organisation;
 - Explanation of the various types of evaluation applied within the organisation (self-evaluation; independent evaluation; centralised/decentralised forms of evaluation; etc);
 - Differentiation of evaluation from other types of assessment carried out within the organisation;
 - Definition of the roles and responsibilities of the evaluation professionals, senior management and programme managers;
 - Emphasis on the need for adherence to the organisation's evaluation guidelines;
 - Explanation of how evaluations are prioritized and planned;
 - Description of how evaluations are organised, managed and budgeted;
 - Emphasis on the management response for the follow-up of evaluations.
6. Policy testing during late 2005 and early 2006 will mainly focus on:
- Defining which evaluations (beyond those that are mandated) that the tracking system for follow-up of evaluation recommendations will cover;
 - Agreeing on the level of programme funds that operational units, in future, will earmark for decentralised evaluation; and
 - The type and coverage of future EO-managed evaluations of the country programmes.
 - Development of a quality assurance system, primarily aimed at supporting the decentralised evaluation system.

Evaluation Planning³

7. Following a canvassing and consultation process that is similar to that found in many other organisations, the Director of the EO has full discretion to select which evaluations will be programmed. These are then identified in a biennial rolling plan, which is posted on the EO's website. The biennial plan of evaluations is not explicitly endorsed

³ *The biennial agenda for 2005-06 can be found at www.undp.org/eo/documents/EvaluationAgenda2005-06.pdf*

by the Executive Board. Evaluations which will be completed or initiated in the coming year are identified in the 'Annual Report of the Administrator on Evaluation' that is presented to the Board each June.

8. Thematic evaluations led by the EO are not, nor have ever been, intended to have full coverage of all UNDP supported activities. The evaluation function, unlike in some organisations, is also not mandated to directly validate the self-assessments of performance against the strategic results framework (the Multi Year Funding Framework) used in UNDP's RBM system. Rather it is intended as a strategic instrument evaluating key corporate, thematic and country issues as determined by the EB, senior management, country counterparts and EO.

Budget and Staffing

Budget

9. The overall budget of the EO in 2005 was approximately USD 5.2 million and has been relatively stable since 2000. The total budget has five major components:

Component 1 (USD 1.5 million annually)

10. This component covers the costs, which are not managed by the EO, of what can be termed as the core staff of the EO, which includes the Director, Deputy Director, two senior Evaluation Advisors, an Analyst, the Operations Manager, and two support staff. This component is assumed to be completely predictable, as long as UNDP continues to operate.

Component 2 (USD 2.5 million annually, but increase to USD 2.9 million for 2006 due to previous underspend)

11. In response to a perception that the EO was underfunded, in 2002, the approach to setting the EO's 'programme' funding, along with that of several other key HQ functions, changed. From 2003, the programme budget ceiling has been set at a fixed amount rather than as a percentage of the UNDP's overall core programme budget. The EO's detailed budget within this ceiling is developed by the EO and then agreed upon by the Bureau of Management and Office of the Administrator before being agreed upon by the Executive Board through the biennial budget approval process. As such, approval of the EO's budget follows general UNDP procedures rather than procedures specific to the EO with any special provisions to secure the independence of the Office.
12. This budget is also considered to be predictable as long as UNDP's core programme budget does not significantly fall. The 'programme' budget is used to fund the remaining professional evaluation positions in the EO (eight positions) which costs approximately USD 1 million per year. The remaining programme budget has been used to fund both establishing and maintaining standards, and knowledge management such as the ERC and EvalNet of the EO and a number of the thematic evaluations requested by the Executive Board.

Component 3 (USD 0.4 million annually)

13. A special programme allocation from UNDP's core programme budget to support the programme of country programme evaluations (termed Assessments of Development Results (ADR) in UNDP). The EO is discussing increasing this budget with senior management, to fund the proposed expansion of coverage of the ADR proposed in the draft Evaluation Policy.

Component 4 (c. USD 0.2 million annually)

14. The Executive Board has requested that when future Global and Regional (five) Framework Programmes are presented to the Board, they be supported by an evaluation of the previous programme. These evaluations are funded by earmarked funds from the budgets of the respective bureaux, which are managed by the EO. These funds are considered to be predictable, since the bureaux are required to provide the necessary funding required by the EO.

Component 5 (c. USD 0.8 million in 2003, 2004 and 2005; to decline to USD 0.4 million from 2006)

15. The EO has 'partnership agreements' with the Dutch, British, Danish, Canadian and Norwegian Governments and has established new agreements within the context of the evaluation of the impact of the Tsunami assistance. These funds have been used to fund the following:
 - Dutch: Evaluation of Governance-Poverty linkages, 4 ADRs, MDG evaluation
 - Danish: Consultation for the 2003 Development Effectiveness Report
 - British: Development of the ADR approach and implementation of c. 6 ADRs
 - Canadian: Gender Mainstreaming Evaluation
 - Norway: Post Conflict evaluation, Gender Mainstreaming Evaluation, Tsunami Evaluation
 - American: Tsunami Evaluation
 - French: Tsunami Evaluation
 - German: Tsunami Evaluation
16. Such funding is predictable during the life of the relevant partnership agreement.
17. The level of funds spent by operational programmes on evaluation is unknown, but based on experience and some analysis of selected programmes in Asia, it is thought, by the EO, to be approximately 3 per cent of the budget.

Staffing

18. The EO is headed by a Director (D2 level which is equivalent to a senior resident representative in a country) who reports directly to the UNDP Administrator. Excluding the Director and Deputy Director, the Office has eleven full-time professional evaluation positions, who work across five functional programme areas.

19. Most professional staff have been recruited from outside of UNDP into the EO, and only five, including the Director and Deputy Director, have been recruited through staff rotation from within UNDP. All EO staff are based in UNDP's New York Headquarters and are contracted on normal UNDP staff contracts.

Products

20. Evaluations completed by the EO between 2002 and 2005 are shown below in Table 2, and are grouped into thematic and country programme evaluations. The evaluations discussed by the Executive Board are shown in bold italics in Table 2.

Table 2: Major products completed by the EO between 2002 and 2005

Year Completed	Thematic Evaluations
2002	i. Post-Conflict Assistance of the Government of Japan through UNDP in Kosovo and East Timor
2003	ii. Evaluation of UNDP's Role in the PRSP process iii. MDG Reports – An Assessment iv. Lessons Learned in Crises and Post-Conflict Situations – The Role of UNDP in Reintegration and Reconstruction Programmes v. Assessment of Micro-macro linkages in poverty alleviation: South Asia
2004	vi. Evaluation of the Second Global Cooperation Framework of UNDP (requested by the Administrator) vii. Institutional Flexibility in Crises And Post-Conflict Situations – Best Practices from the Field
2005	viii. Evaluation of the 2nd regional cooperation framework of the Arab States ix. Evaluation of UNDP's role in the HIV and AIDS Response in Southern Africa (projected completion December 2005)
Year Completed	Assessments of Development Effectiveness (ADRs)
2002	India Sudan
2003	Vietnam Bulgaria
2004	Egypt Jamaica Turkey Ukraine Mozambique
2005	Syria China Honduras (CO comments expected by 17/10/05) Bangladesh (sent for printing) Yemen (being edited)

21. The current pipeline of evaluations that should be completed in 2006/07 is as follows:

Thematics

- Evaluation of Gender Mainstreaming in UNDP (projected completion January 2006)
- Assessment of National Human Development Reports (projected completion March 2006)
- Governance – Poverty Linkages (Joint with the Netherlands)
- Evaluation of UNDP's Role and Contribution in Conflict-Affected Countries
- Evaluation of the 2nd regional cooperation framework of the African States
- Evaluation of the 2nd regional cooperation framework of the Latin American and Caribbean States
- Evaluation of the 2nd regional cooperation framework of the Asia and Pacific States
- Evaluation of the Impact of the Tsunami Response on Local and National Capacities (projected completion April 2006)
- Evaluation of the programme of Technical Cooperation between Developing Countries (TCDC) (*renamed South-South Cooperation*)

Assessment of Development Results

- ADR Rwanda
 - ADR Kenya
 - ADR Mali
 - ADR Serbia and Montenegro
 - ADR Laos
 - ADR Bhutan
 - ADR Paraguay
 - ADR Brazil
 - ADR Jordan
 - ADR Lebanon or Somalia
 - ADR Ethiopia
 - ADR Nigeria
22. The EO launched the Evaluation Resource Centre (ERC) in 2004, which includes outcome and other more strategic evaluations (but not project level) carried out by operational programmes. In response to an Executive Board request, the EO found that the following decentralised evaluations had been produced in 2004:
- Mid-term evaluation of the second regional cooperation framework for Europe and the Commonwealth of Independent States
 - 37 Outcome Evaluations (note that country programme evaluation plans implied that 89 should have been completed in 2004)
 - 165 project evaluations

Evaluation Guidance

23. The EO has produced four sets of Evaluation Guidance. These are:
- The Handbook on Monitoring and Evaluation for Results (2002), which provides guidance to operational staff on monitoring and evaluation
 - ADR Framework Paper (2002), which provides guidance on the purpose, approach and methodology to be used when carrying out an Assessment of Development Results
 - Guidelines for Outcome Evaluators (2002)
 - Guidance on carrying out meta evaluation (2005)
24. The EO is currently planning to update its Guidance for ADRs.

Annex II: The Approach and Methodology for this Assessment

Background and Mandate

In early 2004 the DAC Evaluation Network (sister network of the United Nations Evaluation Group) of which UNDP is an active and longstanding member, endorsed a proposal to test a new approach to evaluating multilateral organisations, primarily to be based on evidence and documentation produced by the multilateral organisations' own monitoring and evaluation systems. The approach builds on the recognition of the growing evaluation capacities of multilateral organisations, and it aims to reduce the many evaluative demands now made on them from other sources, primarily donors. During 2004 a conceptual framework for peer assessment was developed in consultation with UNDP. A working group consisting of Belgium, Canada, Denmark, Germany, Ireland, the Netherlands, Norway and the UK has been established to bring forward the work on the new approach.

This Assessment is guided by the paper "A New Approach to Assessing Multilateral Organisations' Evaluation Performance: Approach and Methodology" agreed upon by participants in the process in June, 2005. This Annex briefly brings together the main lines of the approach pursued, finalized following the meetings and agreements of the peer assessment panel with the UNDP Evaluation Office (EO) in September 2005.

UNDP EO took the lead by volunteering to undergo the first peer assessment of its kind in this area, and indeed proposed the peer approach. Preparation of the exercise with UNDP started in late 2004, and is scheduled to be finalized by the end of 2005.

The approach adopted is based on a "core assessment question," defined and agreed-upon international benchmarks in Norms and relevant practice, and a "peer assessment" mechanism to call upon the judgement of a diverse expert panel with a high level of understanding of these benchmarks and their application.

Once the normative frameworks and process for the assessment were agreed upon with the Evaluation Office (in consultation with the UN Evaluation Group) the succeeding steps were:

Step 2: Collection of data, and its analysis, against these normative frameworks. This took place primarily through extensive documentary research, with a great deal of help from the Evaluation Office, followed by structured and semi-structured interviews with a range of participants and intended users of evaluations. This work was concentrated around five recent evaluations of different types;

Step 3: Agreement by the Panel on the accuracy of the evidence and findings against the frameworks, and on the basis for arriving at judgements;

Step 4: Development of conclusions and recommendations, dialogue on the draft report with the Evaluation Office in a mutual learning process, and finalization of the Panel's report.

Peer Assessment

“Peer review can be described as the systematic examination and assessment of the performance of an entity by counterpart entities, with the ultimate goal of helping the reviewed entity improve its policy making, adopt best practices, and comply with established standards and principles. The examination is conducted on a non-adversarial basis, and it relies heavily on mutual trust among the entities involved in the review, as well as their shared confidence in the process.”^{xix}

Focusing on the core assessment question, the Panel and its advisors planned, assembled and organised the fullest information base practicable, brought together and tested through direct interviews by the Peer Panellists.^{xx} This base of evidence and findings then provided a basis for conclusions drawn primarily from the diverse professional experience and judgements of the Peer Panellists themselves, deliberating as a group on an extensive set of subquestions. The device of professional peer assessment, already well tested in the development field, has been agreed to bring a number of special strengths to this undertaking, as illustrated in international processes such as those of the OECD, WTO and more recently NEPAD. Because of the Panel members’ comparative experience, their judgements are expected to be grounded in a sense of realism, and in the shared need for continuing improvement.

- First, it starts with a shared appreciation of the distinctive challenges of work and evaluation in development cooperation, and the fact that we are all constantly striving to improve.
- Second, it can adapt and apply the most pertinent professional principles, norms and standards in coming to an assessment.
- Third, drawing on experienced professional peers from other development institutions (participating as individuals) will maximize the opportunities for mutual learning, sharing relevant experience and lessons.
- Finally, the resulting assessment should carry particular weight, both internally and externally, for the independence and professional credibility of its results.

Although the Panel works to arrive at consensus judgements where possible, there may also be instances where their judgements are different and a variety of possible suggestions are explored. More generally, while the Panel draws on the experience and knowledge of its members to outline some suggestions, pointers and options for action, it is not the Panel’s role to prescribe any particular model for the organisation being assessed, recognizing that each organisation must ultimately decide for itself, based on its particular circumstances.

It should be stressed that this overall exercise is not itself a formal evaluation. – it is a less comprehensive and in-depth assessment but adheres to a rigorous methodology applying the key principles of evaluation while taking full advantage of the particular benefits of a peer mechanism. It is explicit that “the final conclusions will clearly be a judgement” by the Panel concerned. It has been designed to be targeted, and to minimize additional workload demands on UNDP personnel, in EO and elsewhere. As the first exercise of its kind, it has also put particular emphasis on deriving and feeding back experience, lessons learned and possible recommendations for the future.

The Peer Panel

The Working Group referred to above, in consultation with UNDP's Evaluation Office and with the consent of the DAC Working Party on Development Evaluation, agreed to a Peer Panel comprising representatives of the evaluation departments of Denmark, the Netherlands and UK. These three subsequently selected the UN and independent panel members.

The Peer Panel thus established comprised:

Mary Cole: Independent Evaluation Expert. South Africa. Secretary IDEAS. Formerly Head of the Operations Evaluation Unit of the Development Bank of Southern Africa

Niels Dabelstein: Head, Danida Evaluation Department, Ministry of Foreign Affairs, Denmark. Formerly Chair of DAC Working Party on Development Evaluation (Chair of the Panel)

Tony Faint: Advisor, DFID Evaluation Department, UK, Formerly Director of DFID International Division

Ted Kliet: Acting Deputy Director, Policy and Operations Evaluation Department, Ministry of Foreign Affairs, The Netherlands

Luciano Lavizzari: Director, Evaluation Office, International Fund for Agricultural Development (IFAD)

The team received invaluable assistance from two advisors: *Bernard Wood and Paul Balogun.*

Key Pillars

As agreed in the June 2005 Approach & Methodology, the assessment is centred on the “core assessment question” of:

‘whether a multilateral agency’s own central evaluation office produces evaluations which are credible, valid and useable for learning and accountability purposes as tested by internationally recognised evaluation peers.’

Further, the Approach & Methodology (designed to apply to a wide range of multilateral organisations) hinges on using normative frameworks and **clustering** the treatment of the many issues **under three crucial aspects of evaluation**, specified as:

- A. “**Independence** of evaluations and evaluation systems. “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.”
- B. **Credibility** of evaluations. “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 successes as well as

failures. Recipient countries should, as a rule, fully participate in evaluation in order to promote credibility and commitment.”

- C. **Utility** of evaluations. “To have an impact on decisionmaking, 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 of the many parties involved in development cooperation.”⁴

Normative Framework

It was also agreed in the Approach (Paragraph 18) that assessments should take the new UNEG Norms and Standards (N&S) as the point of departure for UN organisations, to be cross-checked against the relevant issues in the 1998 Review of DAC Principles and other sources suggested. Although these three key aspects are not specifically defined in the N&S we have verified that a composite of the various relevant N&S yields implicit definitions that appear fully consistent with those of the widely accepted Principles, and cover the issues quite fully.

Thus, the Panel has agreed that there was little need to generate yet another framework and possibly cause confusion and disruption and thus that it should use the **UNEG Norms as the guiding framework**, as in the document “Proposed Normative Framework for UNDP Peer Assessment.” The assessment then drew on the **UNEG Standards as instruments for applying the Assessment**, supplemented by other questions judged relevant by the Panel.⁵

For the initial purpose of ensuring proper coverage of the three **main aspects**, the UNEG Norms and Standards were broadly regrouped and checked under those three headings, and as mandated, they would be expected to reemerge as the first issues featured in the Assessment report. At the same time, however, the distinctive coverage and emphases of the new UNEG Norms and Standards, and UNDP’s own particular priorities and ways of working, were obviously prominently reflected in the course of the assessment, and the Panel’s findings and conclusions. The UNEG self-assessment “quality stamp” questionnaire and the self-assessment responses by UNDP have been taken fully into account as very valuable inputs to the assessment design, preparations and information base. UNDP Guidance was also checked and, as soon as it was received, UNDP’s new draft Evaluation Policy was registered under relevant the relevant Norms and main aspects, indicating future directions proposed.

Because of the compressed timeframe for this process and the need to minimize additional workload demands, much necessary preparatory work was carried out in parallel with the processes outlined above in preparing and securing agreement on the Normative framework:

⁴ *These definitions are taken from the OECD DAC Principles for Evaluation of Development Assistance.*

⁵ *Indicators of independence are broadly covered by UNEG Norms N6.1-N6.5 and amplified in the relevant Standards. Indicators of credibility are mainly treated in UNEG Norms N5.1-N5.3, N8.1, N9.1-N9.3 and N11.1-N 11.5 and amplified in the relevant Standards. Indicators of utility are mainly treated in UNEG Norms N2.6, N1.3, N 8.2, N10.1, N 10.2 and N.12.1-N12.3 and amplified in the relevant Standards*

1. Reviewing the first drafting of parts of a preliminary assessment from November 2004, already fact-checked with UNDP EO, to transpose, cross-check, up-date, and fill gaps as needed to reflect the proposed assessment framework and changes at the UNDP end;
2. Extensive review and updating of the documentary and website information base gathered in 2004, supplemented by additional draft and other materials supplied by UNDP EO in mid-October meetings. This base has proved to be extensive and useable, and UNDP EO has continued to offer and provide further materials as required.
3. To make the informational foundations as concrete, focused and systemic as possible within the scope this assessment, a small group of recent evaluation processes and products of different types was selected by the Panel, in close collaboration with UNDP EO, as reference cases taking into account a combination of several criteria.^{6/7} These cases are the GCF II and PRSP evaluations, and Arab States' RCF presented to the Board, and ADRs for Turkey and Mozambique. These were not expected to be full process or product assessments. The reference cases provided a basis for more detailed information gathering by the Panel's advisors, including structured and semi-structured interviews with internal and external participants in these processes from different perspectives. (In total, 21 such interviews were carried out). The evidence was fed into the compilation of findings, with some preliminary triangulation among the various sources and types of information;
4. Advisors provided the Panel with the information base to date, specifying the sources, and organised under Policies to date, Practices, and Future Directions (with the latter reflecting draft policy and innovations now being tested). For practicability, this information was inserted into the existing Proposed Framework. It did not go beyond findings at this stage. This compilation was also provided to UNDP EO for their information and any rapid feedback needed prior to the panel interviews, and later for refinement if necessary in the draft report;
5. Peer interviews: With the benefit of the information assembled, and its individual and collective examination by the panel, including selected primary sources, and any observations to date from UNDP EO, the Panel conducted wider interviews with a number of UNDP Senior Managers and Executive Board members concerned, the reference cases served as one important focus;
6. Integrating the results and insights from these interviews, the Panel completed the triangulation, refinement and confirmation of its base of evidence and findings. On that basis, it then moved into the "judgement phase" with the following steps: panel agreed on its main frameworks for judgements in relation to the Norms and main aspects of the assessment;⁸ the Panel considered and debated the evidence and findings, and arrived at its draft conclusions and recommendations, agreed on draft report;

6 *These featured the need for the examples to be relatively recent, relevant to the main types of UNDP work and evaluations, at different stages of follow-up, and with reasonable accessibility of knowledgeable informants.*

7 *This selection process was completed on 14 October, 2005.*

8 *As a preliminary indication, for example, under Independence criteria would be expected to include such questions as arrangements for the positioning, budgeting and staffing, programming, implementation, reporting, dissemination and follow-up of evaluation.*

7. Draft report served as the basis for the Peer Assessment meeting with UNDP EO to consider the results (held on 30 November, 2005);
8. The Panel carried out final review and any necessary revision to the draft report;
9. The Panel transmits its final report to UNDP EO, providing final opportunity for review and reflecting dissenting views or other key responses;
10. The Panel presents Peer Assessment Report to UNDP and then to the DAC Evaluation Network and UNEG with its report on lessons learned from the exercise.

Annex III: The United Nations Norms for Evaluation

Organised as the Normative Framework for the Peer Assessment

The UN Norms “seek to facilitate system-wide collaboration on evaluation by ensuring that evaluation entities within the UN follow agreed-upon basic principles. They provide a reference for strengthening, professionalizing and improving the quality of evaluation in all entities of the United Nations system.”

The Norms are clustered under three main aspects of evaluation: utility, independence and credibility as the focus of the Peer Assessment. Both the Norms and the three categories are interrelated and overlapping, so that their actual placement here is less important than the fact that they are covered appropriately.

Use of centrally produced evaluations

N 1.1

Purposes of evaluation include understanding why and the extent to which intended and unintended results are achieved, and their impact on stakeholders.

Evaluation is an important source of evidence of the achievement of results and institutional performance.

Evaluation is also an important contributor to building knowledge and to organisational learning.

Evaluation is an important agent of change and plays a critical and credible role in supporting accountability.

N 1.3

Evaluation feeds into management and decisionmaking processes, and makes an essential contribution to managing for results.

Evaluation informs the planning, programming, budgeting, implementation and reporting cycle.

It aims at improving the institutional relevance and the achievement of results, optimizing the use of resources, providing client satisfaction and maximizing the impact of the contribution of the UN system.

N 1.5

Evaluation is not a decisionmaking process per se, but rather serves as an input to provide decisionmakers with knowledge and evidence about performance and good practices.

Although evaluation is used to assess undertakings, it should provide value-added for decision-oriented processes to assist in the improvement of present and future activities, projects, programmes, strategies and policies.

Thus, evaluation contributes to institutional policymaking, development effectiveness and organisational effectiveness.

N 1.7

Evaluation is therefore about *Are we doing the right thing?* It examines the rationale, the justification of the undertaking, makes a reality check and looks at the satisfaction of intended beneficiaries.

Evaluation is also about *Are we doing it right?*

It assesses the effectiveness of achieving expected results.

It examines the efficiency of the use of inputs to yield results.

Finally, evaluation asks *Are there better ways of achieving the results?* Evaluation looks at alternative ways, good practices and lessons learned.

N 2.6

The Governing Bodies and/or Heads of organisations and of the evaluation functions are responsible for ensuring that evaluation contributes to decisionmaking and management. They should ensure that a system is in place for explicit planning for evaluation and for systematic consideration of the findings, conclusions and recommendations contained in evaluations. They should ensure appropriate follow-up measures including an action plan, or equivalent appropriate tools, with clear accountability for the implementation of the approved recommendations.

N 2.7

The Governing Bodies and/or Heads of organisations and of the evaluation functions are responsible for ensuring that there is a repository of evaluations and a mechanism for distilling and disseminating lessons to improve organisational learning and systemic improvement. They should also make evaluation findings available to stakeholders and other organisations of the UN system as well as to the public.

N 4.1 [*This and Norm 4.2 speak clearly to concerns of internal consistency, effectiveness and use*]

Proper application of the evaluation function implies that there is a clear intent to use evaluation findings.

In the context of limited resources, the planning and selection of evaluation work has to be carefully done.

Evaluations must be chosen and undertaken in a timely manner so that they can and do inform decisionmaking with relevant and timely information.

Planning for evaluation must be an explicit part of planning and budgeting of the evaluation function and/or the organisation as a whole. Annual or multi-year evaluation work programmes should be made public.

N 4.2

The evaluation plan can be the result of a cyclical or purposive selection of evaluation topics.

The purpose, nature and scope of evaluation must be clear to evaluators and stakeholders. The plan for conducting each evaluation must ensure due process to ascertain the timely completion of the mandate, and consideration of the most cost-effective way to obtain and analyse the necessary information.

N 10.1 & N 10.2

Transparency and consultation with the major stakeholders are essential features in all stages of the evaluation process. This improves the credibility and quality of the evaluation. It can facilitate consensus building and ownership of the findings, conclusions and recommendations.

Evaluation Terms of Reference and reports should be available to major stakeholders and be public documents. Documentation on evaluations in easily consultable and readable form should also contribute to both transparency and legitimacy.

N 12.1 – 12.3 (Follow-up)

Evaluation requires an explicit response by the governing authorities and management addressed by its recommendations. This may take the form of a management response, action plan and/or agreement clearly stating responsibilities and accountabilities.

There should be a systematic follow-up on the implementation of the evaluation recommendations that have been accepted by management and/or the Governing Bodies.

There should be a periodic report on the status of the implementation of the evaluation recommendations. This report should be presented to the Governing Bodies and/or the Head of the organisation.

N 13.1 – 13.2 (Contribution to Knowledge-building)

Evaluation contributes to knowledge-building and organisational improvement. Evaluations should be conducted and evaluation findings and recommendations presented in a manner that is easily understood by target audiences.

Evaluation findings and lessons drawn from evaluations should be accessible to target audiences in a user-friendly way. A repository of evaluation could be used to distil lessons that contribute to peer learning and the development of structured briefing material for the training of staff. This should be done in a way that facilitates the sharing of learning among stakeholders, including the organisations of the UN system, through a clear dissemination policy and contribution to knowledge networks.

Independence of evaluations and evaluation systems

N 2.1 – 2.4 (Responsibility for Evaluation)

The Governing Bodies and/or the Heads of organisations in the UN system are responsible for fostering an enabling environment for evaluation and ensuring that the role and function of evaluation are clearly stated, reflecting the principles of the UNEG Norms for Evaluation, taking into account the specificities of each organisation's requirements.

The governance structures of evaluation vary. In some cases it rests with the Governing Bodies in others with the Head of the organisation. Responsibility for evaluation should be specified in an evaluation policy.

The Governing Bodies and/or the Heads of organisations are also responsible for ensuring that adequate resources are allocated to enable the evaluation function to operate effectively and with due independence.

The Governing Bodies and/or Heads of organisations and of the evaluation functions are responsible for ensuring that evaluations are conducted in an impartial and independent fashion. They are also responsible for ensuring that evaluators have the freedom to conduct their work without repercussions for career development.

N 6.1 – 6.5 (Independence)

The evaluation function has to be located independently from the other management functions so that it is free from undue influence and that unbiased and transparent reporting is ensured. It needs to have full discretion in submitting directly its reports for consideration at the appropriate level of decisionmaking pertaining to the subject of evaluation.

The Head of evaluation must have the independence to supervise and report on evaluations as well as to track follow-up of management's response resulting from evaluation.

To avoid conflict of interest and undue pressure, evaluators need to be independent, implying that members of an evaluation team must not have been directly responsible for the policymaking, design, or overall management of the subject of evaluation, nor expect to be in the near future.

Evaluators must have no vested interest and have the full freedom to conduct impartially their evaluative work, without potential negative effects on their career development. They must be able to express their opinion in a free manner.

The independence of the evaluation function should not impinge the access that evaluators have to information on the subject of evaluation.

Credibility of evaluations

N 1.2

An evaluation is an assessment, as systematic and impartial as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area, institutional performance, etc. It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof. It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions and contributions of the organisations of the UN system. An evaluation should provide evidencebased information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decisionmaking processes of the organisations of the UN system and its members.

N 1.4

There are other forms of assessment being conducted in the UN system. They vary in purpose and level of analysis, and may overlap to some extent. Evaluation is to be differentiated from appraisal, monitoring, review, inspection, investigation, audit, research, and internal management consulting (all defined in the full text of the Norm).

N 2.5

The Governing Bodies and/or Heads of organisations are responsible for appointing a professionally competent Head of evaluation, who in turn is responsible for ensuring that the function is staffed by professionals competent in the conduct of evaluation.

N 3.1

Each organisation should develop an explicit policy statement on evaluation. The policy should provide a clear explanation of the concept, role and use of evaluation within the organisation, including the institutional framework and definition of roles and responsibilities; an explanation of how the evaluation function and evaluations are planned, managed and budgeted; and a clear statement on disclosure and dissemination.

N 5.1 – 5.3 (Impartiality)

Impartiality is the absence of bias in due process, methodological rigour, consideration and presentation of achievements and challenges. It also implies that the views of all stakeholders are taken into account. In the event that interested parties have different views, these are to be reflected in the evaluation analysis and reporting.

Impartiality increases the credibility of evaluation and reduces the bias in the data gathering, analysis, findings, conclusions and recommendations. Impartiality provides legitimacy to evaluation and reduces the potential for conflict of interest.

The requirement for impartiality exists at all stages of the evaluation process, including the planning of evaluation, the formulation of mandate and scope, the selection of evaluation teams, the conduct of the evaluation and the formulation of findings and recommendations.

N 7.1 – 7.2 (Evaluability)

During the planning stage of an undertaking, evaluation functions can contribute to the process by improving the ability to evaluate the undertaking and by building an evaluation approach into the plan. To safeguard independence this should be performed in an advisory capacity only.

Before undertaking a major evaluation requiring a significant investment of resources, it may be useful to conduct an evaluability exercise. This would consist of verifying if there is clarity in the intent of the subject to be evaluated, sufficient measurable indicators, assessable reliable information sources and no major factor hindering an impartial evaluation process.

N 8.1 (Quality of Evaluation)

Each evaluation should employ design, planning and implementation processes that are inherently quality oriented, covering appropriate methodologies for data collection, analysis and interpretation.

N 8.2

Evaluation reports must present in a complete and balanced way the evidence, findings, conclusions and recommendations. They must be brief and to the point and easy to understand. They must explain the methodology followed, highlight the methodological limitations of the evaluation, key concerns and evidence-based findings, dissident views and consequent conclusions, recommendations and lessons. They must have an executive summary that encapsulates the essence of the information contained in the report, and facilitate dissemination and distillation of lessons.

N 9.1 – 9.3 (Competencies for Evaluation)

Each organisation of the UN system should have formal job descriptions and selection criteria that state the basic professional requirements necessary for an evaluator and evaluation manager. The Head of the evaluation function must have proven competencies in the management of an evaluation function and in the conduct of evaluation studies. Evaluators must have the basic skill set for conducting evaluation studies and managing externally hired evaluators.

N 11.1 – 11.5 (Evaluation Ethics)

Evaluators must have personal and professional integrity.

Evaluators must respect the right of institutions and individuals to provide information in confidence and ensure that sensitive data cannot be traced to its source. Evaluators must take care that those involved in evaluations have a chance to examine the statements attributed to them.

Evaluators must be sensitive to beliefs, manners and customs of the social and cultural environments in which they work. In light of the United Nations Universal Declaration of Human Rights, evaluators must be sensitive to and address issues of discrimination and gender inequality.

Evaluations sometimes uncover evidence of wrongdoing. Such cases must be reported discreetly to the appropriate investigative body. Also, the evaluators are not expected to evaluate the personal performance of individuals and must balance an evaluation of management functions with due consideration for this principle.

Annex IV: A Note on the Draft Evaluation Policy for UNDP

As of September 2005, UNDP has a draft “Evaluation Policy Statement” for the organisation which the Administrator and Senior Management Team have endorsed for testing, with a view to final approval by the Executive Board in 2006. To help make the Panel’s work as relevant and useful as possible, the Evaluation Office made available this draft policy, and the Panel has made extensive references to its content and to possible improvements. As it does not yet have formal standing – not having been presented to UNDP’s Executive Board – the draft policy itself is not included with this report.

Annex V: List of Those Consulted

UNDP

<i>Kemal Dervi</i>	Administrator
<i>Zéphirin Diabré</i>	Associate Administrator
<i>Gilbert Houngbo</i>	Director and Chief of Staff, Office of the Administrator
<i>Saraswathi Menon</i>	Evaluation Office
<i>Nurul Alam</i>	Evaluation Office
<i>Sukai Prom-Jackson</i>	Evaluation Office
<i>Fadzai Gwaradzimba</i>	Evaluation Office
<i>Ruby Sandhu Rojon</i>	Evaluation Office (2004 only)
<i>Suppiramaniam Nanthikesan</i>	Evaluation Office (2004 only)
<i>David Rider Smith</i>	Evaluation Office
<i>Ruth Abraham</i>	Evaluation Office
<i>Khaled Ehsan</i>	Evaluation Office
<i>Anish Pradhan</i>	Evaluation Office
<i>Ravi Rajan</i>	Operations Support Group (2004 only)
<i>Elena Tischenko</i>	Operations Support Group (2004 only)
<i>Vineet Bhatia</i>	Operations Support Group (2004 only)
<i>Abdul Hannan</i>	Programme Specialist, Operations Support Group
<i>Jocelline Bazile-Finley</i>	Deputy Assistant Administrator, Bureau of Management
<i>Romesh Muttukumaru</i>	Bureau of Management (2004 only)
<i>Thomas Eriksson</i>	Office of Corporate Planning (2004 only)
<i>Shoji Nishimoto</i>	Director, Bureau for Development Policy
<i>Walter Franco (retired)</i>	Bureau for Development Policy
<i>Terry McKinley</i>	Bureau for Development Policy
<i>Alvaro Rodriguez</i>	Bureau for Development Policy
<i>Bruce Jenks</i>	Bureau for Resources and Strategic Partnerships
<i>Nicola Harrington</i>	Bureau for Resources and Strategic Partnerships (2004 only)
<i>Stéphane Vigie</i>	Bureau for Resources and Strategic Partnerships (2004 only)
<i>Jessie Mabutas</i>	Director, Office of Audit and Performance Review
<i>Diane Kepler</i>	Internal Audit Section (2004 only)
<i>Neal Walker</i>	Regional Bureau for Latin America (2004 only)
<i>Enrique Ganuza</i>	Regional Bureau for Latin America
<i>Nada Al-Nashif</i>	Chief, Regional Programme Division, Regional Bureau of Arab States
<i>Walid Badawi</i>	Regional Bureau of Arab States
<i>Azza Karam</i>	Regional Bureau of Arab States

<i>Marta Ruedas</i>	Deputy Regional Director, Regional Bureau for Europe and the Commonwealth of Independent States
<i>Claire Van der Vaeren</i>	former Deputy Resident Representative, Turkey
<i>Elizabeth Lwanga</i>	Deputy Regional Director, Southern and Eastern Africa, Regional Bureau for Africa
<i>Jeffrey Avina</i>	Deputy Regional Director, Western Central Africa, Regional Bureau for Africa
<i>Jacques Loup</i>	Regional Bureau for Africa (2004 only)
<i>Nardos Bekele-Thomas</i>	Regional Bureau for Africa (2004 only)
<i>Shigeki Komatsubara</i>	Regional Bureau for Africa (2004 only)
<i>Sam Barnes</i>	Senior Advisor, Bureau Crisis Prevention and Recovery
<i>Marylene Spezzati</i>	Resident Representative, Mozambique, and colleagues in the Country Office

Consultants on EO evaluations

Ayse Ayata
Michael Bamberger
Richard Flaman
James Freedman
Rajeev Pillay
Michael Reynolds
Johannes Linn
John Weeks

UNDGO

Sally Fegan-Wyles Director, Development Group Office
Lilana Ramirez

Canada – UN Mission in New York

Diana Rivington Counsellor (Development)

Denmark – UN Mission in New York

Pia Starbæk Szczepanski (2004 only)

Mozambique – UN Mission in New York

Antonie Macheve

Norway – UN Mission in New York

Kjersti Rodsmoen (2004 only)

The Netherlands – UN Mission in New York

Kitty van der Heijden (2004 only)

Turkey – UN Mission in New York

Damla Say Turkish Representative

UK – UN Mission in New York

Phil Evans (2004 only)

Representatives from all countries serving on the Executive Board were invited to take part in an informal consultation with Panel Members on the morning of 3rd November 2005. Participants from the following countries were able to take part: Austria, Australia, Canada, China, Denmark, El Salvador, Eritrea (Chair of the meeting), France, Gambia, Germany, Italy, Japan, the Netherlands, New Zealand, Norway, UK, and the USA.

Annex VI: Compilation of Evidence and Findings

Very comprehensive evidence and findings against the selected UNEG Norms and Standards were compiled during the Panel's work. Judgements made in the Panel's report use this data but are also based on the many interviews conducted by the Panel which are not included to protect the anonymity of interviewees. The document summarizing this compilation (90 pages) is available on request from UNDP Evaluation Office: *contact.eo@undp.org*

and from Danida Evaluation Department: *eval@um.dk*

Endnotes

- i See Annex II for more details of the development of this approach and its first application in this case.
- ii See Annex II for more details of the development of this approach and its first application in this case.
- iii As of September 2005, UNDP has a draft “Evaluation Policy Statement” for the organisation which the Administrator and Senior Management Team have endorsed for testing, with a view to final approval by the Executive Board in 2006. To help make the Panel’s work as relevant and useful as possible, the Evaluation Office made available this draft policy, and the Panel has made extensive references to its content and to possible improvements. As it does not yet have formal standing – not having been presented to UNDP’s Executive Board – the draft policy itself is not included with this report.
- iv See Annex II for more details of the development of this approach and its first application in this case.
- v These are amplified as follows:
“Human Development and Human Rights: UNDP advocates a people-centred approach to development that enhances capabilities, choices and rights for all men and women. Its support to countries’ development goals is inspired by universally shared values of equity, justice, gender equality and respect for diversity. Evaluation in UNDP is also guided by its vision of people-centred, gender-responsive and participatory development that reduces socio-economic disparities.
UN Coordination and Global Partnership: Evaluation draws on and contributes to collaboration within the UN system to improve effectiveness and reduce transaction costs for development cooperation. UNDP advocates strengthened professional collaboration under the aegis of UNEG and country-level coordination in evaluation under the resident coordinator system. With the increasing engagement of UNDP in global initiatives and partnership programs with other donors, NGOs and civil society, joint evaluations enhance global partnership as well as partnercountry capacity for evaluation.”
- vi Adapted from “Peer Review: a tool for cooperation and change: An Analysis of an OECD Working Method, OECD, 2002.”
- vii The Peer Panel for the assessment of UNDP consisted of three senior evaluators from bilateral donors – Denmark, the Netherlands and the UK; a sister multilateral agency, IFAD; and an international evaluation expert from an UNDP programme country, South Africa. The Peer Panel’s work was supported by two advisors.
- viii For a more detailed description, see Annex 1.
- ix DAC, OECD, Principles for Evaluation of Development Assistance.

- x See, “Evaluating Country Programmes, Vienna Workshop, 1999”, Evaluation and Aid Effectiveness series # 2, OECD (1999).
- xi DAC, OECD, Principles for Evaluation of Development Assistance.
- xii The differences between evaluation and other forms of assessment are discussed in detail in UNDP’s Handbook of Monitoring and Evaluation (2002).
- xiii The exception is the Tsunami Evaluation which is more akin to an inventory of actions, rather than an evaluation in the conventional sense.
- xiv These evaluations were: (i) Evaluation of UNDP’s Role in the PRSP process; (ii) MDG Reports – An Assessment; (iii) Evaluation of the Second Global Cooperation Framework of UNDP; (iv) Evaluation of the 2nd Regional Cooperation Framework for the Arab States; (v) Evaluation of UNDP’s role in the HIV and AIDS Response in Southern Africa; and (vi) Evaluation of Gender Mainstreaming in UNDP.
- xv Evaluation of UNDP’s Regional Cooperation Framework for the Arab States (2002-2005), UNDP, 2005.
- xvi Jamaica ADR, UNDP, 2005.
- xvii DAC, OECD, Principles for Evaluation of Development Assistance.
- xviii See “An assessment of UNDP practices in monitoring and evaluation at the country level” (2005) and “Are Outcome Evaluations Working?” A review of outcome evaluation methodology and practice in UNDP.” (2004)
- xix Adapted from “Peer Review: a tool for cooperation and change: An Analysis of an OECD Working Method, OECD, 2002.”
- xx The Peer Panel for the assessment of UNDP consisted of three senior evaluators from bilateral donors – Denmark, the Netherlands and the UK; a sister multilateral agency, IFAD; and an international evaluation expert from an UNDP programme country, South Africa. The Peer Panel’s work was supported by two advisors.

Who evaluates the evaluator? Occasionally the independence, credibility and utility of evaluation are challenged both by those being evaluated and by the users of evaluation.

This assessment is the first of a new approach, designed under the auspices of the OECD/DAC Network on Development Evaluation. It aims at assessing and enhancing multilateral agencies' own evaluation capacity and performance, with an ultimate view to improving their development performance. The approach used is a "peer assessment". UNDP volunteered to undergo the first such assessment.

Three crucial aspects of evaluation – credibility, independence, and utility – were adopted as the broad yardsticks in assessment. The Norms for Evaluation in the UN System, formally adopted by the UN Evaluation Group were used as benchmarks. The Peer Panel has made its assessment against these benchmarks as well as against prevailing international practise.

The overall conclusion of the Peer Panel is that UNDP's Evaluation Office enjoys an acceptable level of independence and produces evaluations that are credible, valid and useful for learning and strategy formation in the organisation. At the same time, it's potential for strengthening accountability and performance assessment is being under-exploited, both for the purpose of accountability and as an essential basis for learning.

Danida



**Buitenlandse
Zaken**

